RL-TR-95-47, Vol II (of two)
In-House Report
February 1995

PROCEEDINGS OF THE 1994 ANTENNA APPLICATIONS SYMPOSIUM

Daniel H. Schaubert, et al

DTIC
SELECTED
APR 0 7 1995
G

APPROVED FOR PUBLIC RELEASE; DISTRIBUTION UNLIMITED.

19950406 034

**Rome Laboratory
Air Force Materiel Command
Griffiss Air Force Base, New York**

This report has been reviewed by the Rome Laboratory Public Affairs Office (PA) and is releasable to the National Technical Information Service (NTIS). At NTIS it will be releasable to the general public, including foreign nations.
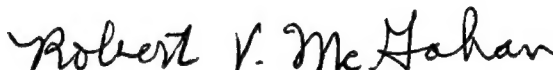
RL-TR-95-47, Vol II (of two) has been reviewed and is approved for publication.

APPROVED: *Dan Jacavanco*

DANIEL J. JACAVANCO, Chief
Antennas & Components Division
Electromagnetics & Reliability Directorate

FOR THE COMMANDER: *Robert V. McGahan*

ROBERT V. McGAHAN
Acting Director
Electromagnetics & Reliability Directorate

If your address has changed or if you wish to be removed from the Rome Laboratory mailing list, or if the addressee is no longer employed by your organization, please notify RL (ERAA) Hanscom AFB MA 01731. This will assist us in maintaining a current mailing list.

Do not return copies of this report unless contractual obligations or notices on a specific document require that it be returned.

REPORT DOCUMENTATION PAGE

Form Approved
OMB No. 0704-0188

1. AGENCY USE ONLY (Leave Blank)	2. REPORT DATE February 1995	3. REPORT TYPE AND DATES COVERED In-House

4. TITLE AND SUBTITLE
PROCEEDINGS OF THE 1994 ANTENNA APPLICATIONS SYMPOSIUM

6. AUTHOR(S)
Daniel H. Schaubert, et al

5. FUNDING NUMBERS
PE - 62702F
PR - 4600
TA - 14
WU - PW

7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES)
Rome Laboratory (ERAA)
31 Grenier Street
Hanscom AFB MA 01731-3010

8. PERFORMING ORGANIZATION REPORT NUMBER
RL-TR-95-47, Vol II (of two)

9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES)
Rome Laboratory (ERAA)
31 Grenier Street
Hanscom AFB MA 01731-3010

10. SPONSORING/MONITORING AGENCY REPORT NUMBER

11. SUPPLEMENTARY NOTES Rome Laboratory Project Engineer: Zachary O. White/ERAA (617) 478-3191. Volume I consists of pages 1 through 258. Volume II consists of pages 259 through 500.

12a. DISTRIBUTION/AVAILABILITY STATEMENT
Approved for public release; distribution unlimited.

12b. DISTRIBUTION CODE

13. ABSTRACT (Maximum 200 words)
The Proceedings of the 1994 Antenna Applications Symposium is a collection of state-of-the-art papers relating to phased array antennas, multibeam antennas, satellite antennas, microstrip antennas, reflector antennas, HF, VHF, UHF and various other antennas.

14. SUBJECT TERMS Antennas, Satellite Antennas, Broadband Antennas, Microstrip, Reflector, HF, VHF, UHF, Multibeam Antennas, Array Antennas

15. NUMBER OF PAGES 248

16. PRICE CODE

17. SECURITY CLASSIFICATION OF REPORT UNCLASSIFIED	18. SECURITY CLASSIFICATION OF THIS PAGE UNCLASSIFIED	19. SECURITY CLASSIFICATION OF ABSTRACT UNCLASSIFIED	20. LIMITATION OF ABSTRACT SAR

CONTENTS

III. Analysis and Measurements

IV. Arrays and Feeds

V.   ELEMENTS

Accesion For		
NTIS CRA&I	☒	
DTIC TAB	☐	
Unannounced	☐	
Justification		
By		
Distribution /		
Availability Codes		
Dist	Avail and / or Special	
A-1		

A FINITE ELEMENT CODE FOR
PHASED ARRAY RADIATOR ANALYSIS

Daniel T. McGrath

USAF Phillips Laboratory, PL/WSR
3550 Aberdeen Ave. SE
Kirtland AFB, NM 87117-5776

1. INTRODUCTION

Successful design of radiators for large phased array antennas has usually required the use of accurate numerical methods for predicting the scan-dependent gain and impedance [1],[2]. Yet, two essential problems were apparent: First, attempts to develop phased arrays for broadband applications led to radiator designs whose structures could not be adequately modeled with existing numerical methods; and second, computer codes were fairly specialized, each limited to a fairly narrow class of geometries. The need for a method that could account for radiators with arbitrary geometry, including complex feed structures, irregular conductor configurations and inhomogeneous dielectrics led to consideration of the finite element method (FEM). That method has been used successfully for similar problems in electromagnetic scattering by using integral equations to impose radiation conditions [3]-[5]. Applying the technique to the phased array problem required development of similar boundary conditions for periodic structures and waveguide feeds.

The analysis and results described in this paper consider the general case of one radiator in an infinite array, fed through a ground plane by a cylindrical waveguide. Accurate feed modeling is accomplished by enforcing continuity, across the ground plane

aperture, between the FEM solution and an arbitrary number of waveguide modes. A periodic integral equation is imposed at a plane above the antenna's physical structure to enforce the radiation condition and to confine the analysis to a single array unit cell. The electric field is expanded in terms of vector finite elements inside the unit cell, and Galerkin's method is used to write the problem as a matrix equation. The Floquet conditions on the unit cell side walls are enforced as transformations of the matrix that effectively "wrap" opposing unit cell walls onto each other with a phase shift appropriate to the scan angle and unit cell dimensions. The solution of the linear system gives a piecewise-linear approximation to the electric field, from which active reflection coefficient and active element gain are computed. Validation tests with a general-purpose computer code confirm that the method is versatile, accurate and reliable.

2. THEORY & FORMULATION

2.1 General Description

Figure 1 illustrates the generic problem geometry and notation conventions. The array is planar, periodic in both the x and y directions, and has an arbitrary skew angle, γ. The volume region Ω is a single unit cell truncated below at the ground plane and waveguide aperture, and at a plane $z=h$ above the radiator's physical structure. The surface Γ bounding Ω is the union of: (a) all perfectly conducting surfaces; (b) the "radiation boundary," Γ_R; (c) the waveguide aperture, Γ_W; and (d) the unit cell side walls Γ_{x+}, Γ_{x-}, Γ_{y+}, and Γ_{y-}. Ω may contain any number of material regions with distinct complex permittivity, ϵ, and permeability, μ, as well as any number of perfectly conducting obstacles, including infinitely thin wires and plates.
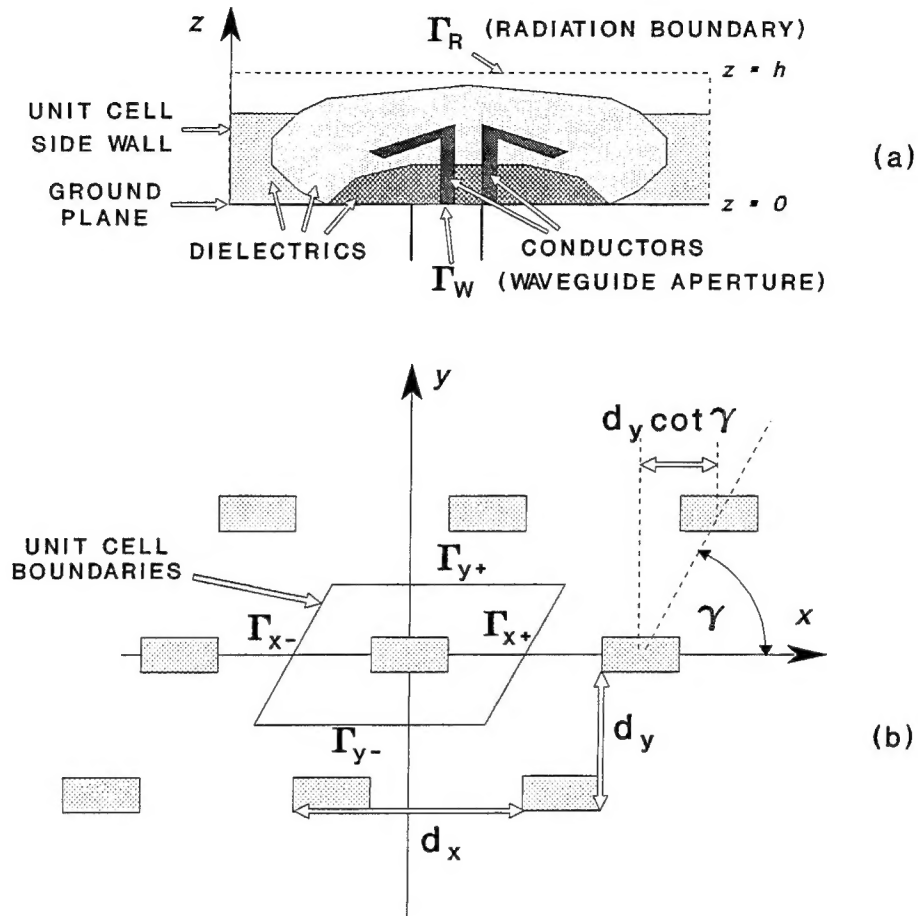
Figure 1. Generic Problem Geometry:
(a) Unit Cell Cross Section; (b) Array Lattice

A numerical description of the problem geometry is created by subdividing Ω into small volume elements, or "cells." In this work, the cells are tetrahedra. The example exploded mesh shown in Figure 2 is for a unit cell of a printed flared notch array. The cells in the center represent the dielectric substrate, while those on the left and right are free space. This radiator is one of the types that motivated this work because it includes an inhomogeneous dielectric and irregularly-shaped conductors.

261
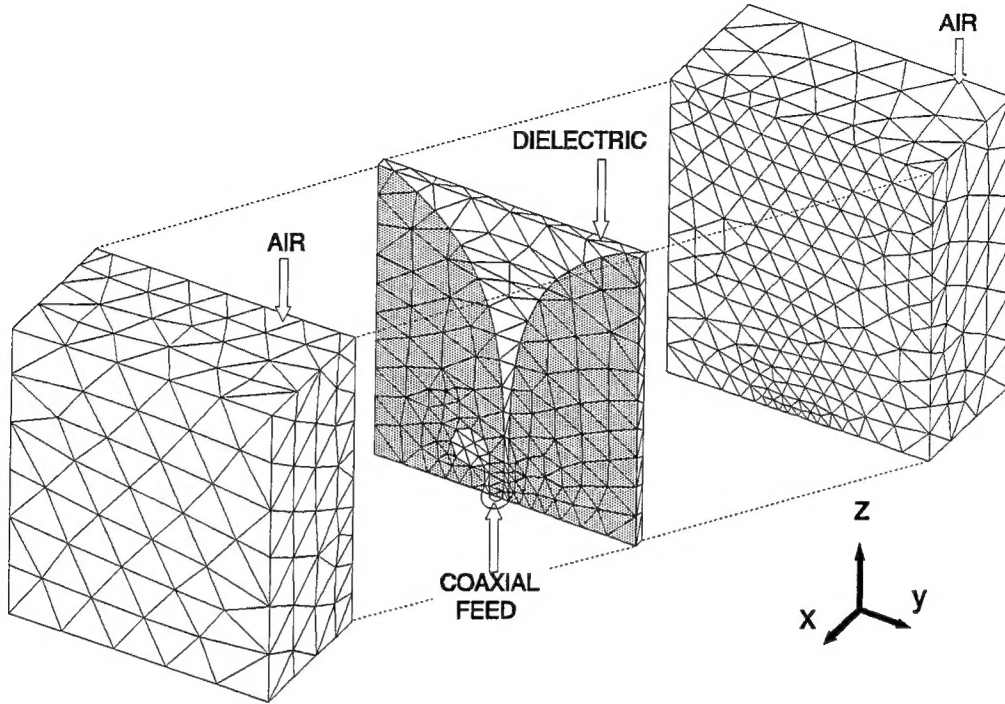
Figure 2. Example Unit Cell Finite Element Mesh (Printed Flared Notch)

2.2 Expansion, Weighting and the Matrix Equation

The electric field everywhere inside Ω and on its bounding surfaces is approximated by the following expansion:

$$\bar{E} \approx \tilde{E} = \sum_{s=1}^{N} e_s \bar{\psi}_s(x,y,z) \tag{1}$$

where each e_s is a complex scalar and each $\bar{\psi}_s$ is a real-valued vector function. N is the number of mesh edges that are not on perfectly conducting surfaces. Each $\bar{\psi}_s$ is defined within the mesh cells adjacent to edge s, and is zero elsewhere. $\bar{\psi}_s$ is the "vector edge element" [4],[6]:

$$\bar{\psi}_s = \ell_{ij}\left(f_i \nabla f_j - f_j \nabla f_i\right) \tag{2}$$

where i and j are the indices of the nodes bounding edge s, and f_i and f_j are the corresponding linear scalar node-based elements [7]. The scaling by the edge length ℓ_{ij} en-

sures that the component of $\bar{\psi}_s$ tangent to the edge is a unit vector. These expansion functions have two important advantages over scalar functions: (a) they are divergence-less, so that the expansion (1) will satisfy the divergence condition; and (b) they enforce continuity of the tangential field components across cell boundaries while allowing the normal component to be discontinuous, so that they correctly model dielectric discontinu-ities [6],[8].

The electric field must satisfy the source-free vector wave equation

$$L(\bar{E}) = \nabla \times \frac{1}{\mu_r} \nabla \times \bar{E} - k_0^2 \epsilon_r \bar{E} = 0 \tag{3}$$

where k_0 is the free space wavenumber. This form allows ϵ_r and μ_r to be functions of position. The residual error due to the approximation (1) is $\bar{R} = L(\bar{E})\text{-}L(\tilde{E})$. The weighted residuals procedure sets $<\bar{R}, \bar{W}> = 0$, where $\bar{W}$ is a weighting function and the inner product integration is taken over all of Ω. Applying a Green's identity to the integral produces the "weak form":

$$F(\bar{E}) = \int_\Omega \left[\frac{1}{\mu_r} \nabla \times \bar{W}^* \cdot \nabla \times \bar{E} - k_0^2 \epsilon_r \bar{W}^* \cdot \bar{E} \right] dv$$

$$- jk_0 \eta_0 \int_\Gamma \bar{W}^* \cdot (\hat{n} \times \bar{H}) ds = 0 \tag{4}$$

where $\hat{n}$ is the outward normal to Ω and η_0 is the impedance of free space. (The term *weak form* stems from the fact that the Green's identity shifted one derivative from $\bar{E}$ to $\bar{W}$, weakening the differentiability requirement on $\bar{E}$. This is important here in allowing the use of linear functions to expand $\bar{E}$.) The discretization of this functional using Galerkin's method generates a system of equations. That system is modified by a trans-

formation R representing side wall periodicity conditions, giving:

$$[R][S^I + S^W + S^R][R]^H E = E^{inc} \qquad (5)$$

(Superscript H denotes Hermitian, or conjugate transpose.) The matrix S^I originates in the volume integral in (4), while S^W and S^R originate in the surface integral taken over Γ_W and Γ_R, respectively. The right side vector E^{inc} results from integrating the incident field term over Γ_W. Finally, the solution vector E gives the coefficients e_s in (1).

2.3 Interior Discretization

Initially, the unit cell mesh is treated as though all nonconducting portions of Γ are open-circuit walls, where the mesh is simply truncated without regard to boundary conditions. When the expansion (1) is substituted into (4), it gives one equation in N unknowns. Using Galerkin's method, each $\bar{\psi}_t$ is substituted in turn for $\bar{W}$, giving N equations. The result is that the terms of S^I are given by the volume integrals

$$S^I_{st} = \int_{\Omega_{st}} \left[\mu_r^{-1} \nabla \times \bar{\psi}_s \cdot \nabla \times \bar{\psi}_t - k_0^2 \epsilon_r \bar{\psi}_s \cdot \bar{\psi}_t \right] dv \qquad (6)$$

The domain of integration, Ω_{st}, is the union of all cells shared by edges s and t.

2.4 Waveguide Integral Equation

Satisfying field continuity across Γ_W requires an integral equation for tangential $\bar{H}$ in terms of $\bar{E}$ on Γ. This is obtained in a manner similar to that used by Harrington and Mautz to solve for waveguide radiation with the method of moments [9], starting with a modal expansion from Marcuvitz [10]:

$$\bar{E}^t = e^{-\gamma_0 z} \bar{g}_\nu + \sum_{i=0}^{\infty} C_i e^{\gamma_i z} \bar{g}_i \tag{7}$$

The superscript t denotes transverse (to z), $\bar{g}_i$ is an orthonormal mode function with propagation constant γ_i, and the index i includes all TE_{mn} and TM_{mn} modes. The first term is due to a unit-amplitude incident field in the mode ν, usually the dominant mode. The unknown coefficients C_i may be found from the transverse electric field over Γ_W as

$$C_i = \int_{\Gamma_W} \bar{E}^t \cdot \bar{g}_i \, ds - \delta_{\nu i} \tag{8}$$

where $\delta_{\nu i}$ is a Kronecker delta. C_0 is the reflection coefficient for the dominant waveguide mode. The transverse magnetic field in terms of the modal admittances Y_i is:

$$\bar{H}^t = Y_0 e^{-\gamma_0 z} (\hat{z} \times \bar{g}_0) - \sum_{i=0}^{\infty} C_i Y_i e^{\gamma_i z} (\hat{z} \times \bar{g}_i) \tag{9}$$

Substituting (8) into (9) gives $\bar{H}^t$ in terms of $\bar{E}^t$ on Γ_W. Then evaluating $\hat{n} \times \bar{H}^t$ on Γ_W ($\hat{n} = \hat{z}$) gives the integral equation

$$\hat{n} \times \bar{H}^t \bigg|_{\Gamma_W} = 2 Y_0 \bar{g}_0 - \sum_{i=0}^{\infty} Y_i \bar{g}_i \int_{\Gamma_W} \bar{E}^t \cdot \bar{g}_i \, ds \tag{10}$$

The result of substituting the right side into (4), followed by expansion and testing is

$$S_{st}^W = j k_0 \eta_0 \sum_{i=0}^{\infty} Y_i \Psi_{si} \Psi_{ti} , \quad s, t \in \Gamma_W \tag{11}$$

$$\Psi_{si} = \int_{\Gamma_W} \bar{\psi}_s \cdot \bar{g}_i \, ds \tag{12}$$

$$E_t^{inc} = 2jk_0\eta_0 Y_0 \Psi_{t0}, \ t \in \Gamma_W \tag{13}$$

The inner products of expansion functions and waveguide mode functions may be carried

out in closed form for rectangular waveguide; or numerically for circular and circular

coaxial waveguide [11].

2.5 Periodic Integral Equation

Each radiator in the infinite array is assumed to be excited by a unit-amplitude

incident field in the waveguide from $z<0$. The excitation phases produce a beam direct-

ed toward θ_0,ϕ_0 in spherical coordinates. The fields above $z=0$ must obey the condition:

$$\bar{E}(x+md_x+nd_y\cot\gamma, \ y+nd_y) = \bar{E}(x,y)\,e^{-j\beta_x(md_x+nd_y\cot\gamma)}e^{-j\beta_y nd_y} \tag{14}$$

$$\beta_x = k_0 \sin\theta_0 \cos\phi_0 \tag{15}$$

$$\beta_y = k_0 \sin\theta_0 \sin\phi_0 \tag{16}$$

An integral equation for the magnetic field transverse to Γ_R in terms of the electric field

transverse to Γ_R is [12]:

$$-\hat{n}\times\bar{H}(\Gamma_R) = \sum_{m=-\infty}^{\infty}\sum_{n=-\infty}^{\infty}\ddot{T}_{mn}\cdot\bar{E}_{uc}(k_{xmn},k_{ymn})\,e^{-jk_{xmn}x}e^{-jk_{ymn}y} \tag{17}$$

$$k_{xmn} = k_0\sin\theta_0\cos\phi_0 - \frac{2\pi m}{d_x} \tag{18}$$

$$k_{ymn} = k_0\sin\theta_0\sin\phi_0 - \frac{2\pi n}{d_y} + \frac{2\pi m\cot\gamma}{d_x} \tag{19}$$

The negative imaginary result is chosen when the square root is negative, and $\bar{E}_{uc}(k_x,k_y)$

is the two-dimensional Fourier transform of the unit cell electric field transverse to Γ_R:

$$\ddot{\overline{T}}_{mn} = \frac{\left[k_0^2 - k_{xmn}^2 - k_{ymn}^2\right]^{-1/2}}{2d_x d_y k_0 \eta_0} \begin{bmatrix} (k_0^2 - k_{ymn}^2) & k_{xmn} k_{ymn} \\ k_{xmn} k_{ymn} & (k_0^2 - k_{xmn}^2) \end{bmatrix} \qquad (20)$$

$$\overline{E}_{uc} = \sum_{s=1}^{N} e_s \overline{\xi}_s(k_x,k_y) = \sum_{s=1}^{N} e_s \int_{\Gamma_R} \overline{\psi}_s e^{jk_x x} e^{jk_y y} \, dx \, dy \qquad (21)$$

Substituting this expression into the surface integral of (4), then testing with each $\overline{\psi}_t$ gives

$$S_{st}^R = \sum_m \sum_n \overline{\xi}_{smn} \cdot \ddot{\overline{T}}_{mn} \cdot \int_{\Gamma_R} \overline{\psi}_t e^{-jk_{xmn}x} e^{-jk_{ymn}y} \, dx \, dy \ , \ s,t \in \Gamma_R \qquad (22)$$

where $\overline{\xi}_{smn}$ denotes $\overline{\xi}_s(k_{xmn},k_{ymn})$. The integral in (22) is equivalent to $\overline{\xi}_{tmn}^*$ ($\overline{\psi}_t$ is a real function, therefore its Fourier transform is Hermitian [13:193], i.e. $\overline{\xi}_t(-k_x,-k_y) = \overline{\xi}_t^*(k_x,k_y)$), so the final expression for the entries in the matrix S^R is

$$S_{st}^R = \sum_m \sum_n \overline{\xi}_{smn} \cdot \ddot{\overline{T}}_{mn} \cdot \overline{\xi}_{tmn}^* \ , \ s,t \in \Gamma_R \qquad (23)$$

The edges that are on both Γ_R and a unit cell boundary are treated differently. There are no entries in S^R for edges on Γ_{x+} and Γ_{y+}. Instead, the entries for edges on Γ_{x-} and Γ_{y-} are computed as if the mesh extended into the adjacent unit cells. This is consistent with Gedney's concept of "overlap elements" [14]. It ensures that there is no conflict between the periodic radiation condition and the side-wall periodicity conditions.

2.6 Side-Wall Periodicity Conditions

The mesh is assumed to have been created so that the surface mesh on opposing side walls is identical. Then each mesh edge on Γ_{x+} or Γ_{y+} has an "image edge" (the term does not relate to electric or magnetic images) on the opposing boundary that is

identical except for a translation of $(-d_x,0)$ or $(-d_y\cot\gamma,-d_y)$, respectively. The implementation of Floquet's condition on these boundaries may be viewed conceptually as follows: The Γ_{x+} and Γ_{y+} edges will be merged with their $-x$ and $-y$ counterparts, effectively recreating an infinite mesh by wrapping opposing side walls onto each other. Suppose, for example, that edge #19 is on Γ_{x+} and its image edge on Γ_{x-} is edge #2. The original matrix will have a pair of entries $S_{19,k}$ and $S_{k,19}$ for every edge k that shares at least one cell with edge #19. In transforming the matrix, these entries will be eliminated and new ones added. The new entries are: $S'_{2,k} = S_{19,k}\exp\{j\psi_x d_x\}$ and $S'_{k,2} = S_{k,19}\exp\{-j\psi_x d_x\}$ for edges k that are not also on Γ_{x+}; and $S'_{2,j} = S_{2,j} + S_{19,k}$ and $S'_{j,2} = S_{j,2} + S_{k,19}$ for each edge k on Γ_{x+} whose image edge is edge j. Γ_{y+} is handled similarly, except that the phase shift terms are $\exp\{\pm j(\psi_x d_y\cot\gamma + \psi_y d_y)\}$.

This procedure may be regarded as the transformation

$$S' = R\,S\,R^H \tag{24}$$

The matrix R is $M{\times}N$, where $N{-}M$ is the number of edges on Γ_{x+} and Γ_{y+} combined. Its first M columns are the $M{\times}M$ identity matrix. The remaining columns have an entry R_{ik} for each pair of image edges (ordering edges so that those in Γ_{x+} and Γ_{y+} are last):

$$R_{ik} = \begin{cases} e^{j\beta_x d_x}, & k\in\Gamma_{x+},\, i\in\Gamma_{x-} \\[4pt] e^{j(\beta_x d_y\cot\gamma + \beta_y d_y)}, & k\in\Gamma_{y+},\, i\in\Gamma_{y-} \\[4pt] e^{j(\beta_x d_x + \beta_y d_y\cot\gamma + \beta_y d_y)}, & k\in\Gamma_{x+}\cap\Gamma_{y+},\, i\in\Gamma_{x-}\cap\Gamma_{y-} \end{cases} \tag{25}$$

The transformation does not affect S^W or S^R because those matrices do not have any nonzero entries for edges in Γ_{x+} or Γ_{y+}.

2.7 Reflection Coefficient and Element Gain

The solution of the linear system gives the ve   of coefficient: for pc

ing to fields along the mesh edges $s = 1,2,...M$. The coefficients for the Γ_{x+} and Γ_{y+}

boundaries are then determined using the Floquet conditions.

The active reflection coefficient, R, is determined from those coefficients e_s that

correspond to edges in Γ_W:

$$R = \sum_{s \in \Gamma_W} e_s \Psi_{s0} - 1 \qquad (26)$$

using Ψ_{s0} from (12).

The active element gain is analogous to a transmission coefficient. It is a mea-

sure of the excitation strength of a plane wave (a Floquet mode) propagating away from

the array. Amitay et. al. shows that the θ and ϕ element gain patterns are due to the

lowest order TE and TM Floquet modes, respectively [2:57]. Their expressions in terms

of (21), the Fourier transform of the unit cell field transverse to Γ_R are:

$$E_\theta = \frac{\sec \theta}{\sqrt{d_x d_y}} (\hat{x} \cos\phi + \hat{y} \sin\phi) \cdot \bar{E}_{uc}^*(k_{x00}, k_{y00}) \qquad (27)$$

$$E_\phi = \frac{1}{\sqrt{d_x d_y}} (\hat{y} \cos\phi - \hat{x} \sin\phi) \cdot \bar{E}_{uc}^*(k_{x00}, k_{y00}) \qquad (28)$$

A check for conservation of power may be made by computing the transmission

coefficients for the two mn=00 Floquet modes, whose admittances are Y_{100} and Y_{200} for

TM and TE:

269

$$T_\theta = E_\theta \sqrt{Y_{200}/Y_0} = E_\theta \sqrt{\sec\theta/Y_0\eta_0} \qquad (29)$$

$$T_\phi = E_\phi \sqrt{Y_{100}/Y_0} = E_\phi \sqrt{\cos\theta/Y_0\eta_0} \qquad (30)$$

Y_0 is the feed waveguide's dominant mode admittance. When the feed waveguide supports only one mode and there are no grating lobes in visible space, the conservation of power relationship is $|R_a|^2 + |T_\theta|^2 + |T_\phi|^2 = 1$.

3. VALIDATION AND DEMONSTRATION

The preceding theory was implemented in a general-purpose computer code. The code is geometry-independent, owing to the use of commercial software for geometry definition and mesh generation. It uses the conjugate gradient method and biconjugate gradient methods (CGM) [15] to solve the system of equations. The example results that follow are both validation (comparisons to solutions obtained by other methods) and demonstration (calculations for problems that have not been addressed by other methods). In most cases, the incident field was in the dominant waveguide mode and 32 waveguide modes $(0 \leq m, n \leq 3)$ and 121 Floquet modes $(-5 \leq m, n \leq 5)$ were included in calculating S^W and S^R, respectively.

3.1 Waveguide Arrays

Arrays of open-ended waveguides, horns, and slots are a simplified case of the preceding analysis [16]. These cavity-type elements couple only through apertures in a common, conducting ground plane. Hence, the formulation may be simplified by excluding the matrix transformation (24).

The first example case is an array of open-ended circular waveguides, whose apertures are flush with a conducting ground plane. They may also include dielectric plugs that fill the waveguide to a distance h under the ground plane. The finite element model consists of a circular cylinder of height h, that is, a short section of waveguide. A conducting boundary condition is applied at its outer wall. The waveguide boundary condition is applied at the lower end. The periodic radiation condition is applied at the top end. One subtlety in the modeling of this object is that the size of the model must be adjusted so that the mesh of flat-sided tetrahedra has the same volume as the cylinder being modeled.

Amitay et. al. computed the active reflection coefficient of circular waveguide arrays using the "multimode" method [2:293]. It is an integral equation method that uses the waveguide and Floquet modes as expansion and testing functions. Figure 3 shows those results as continuous curves and the corresponding HFEM results as discrete points. The agreement between the two methods is exceptional, demonstrating the validity of the circular waveguide and periodic boundary conditions used by the HFEM.

The second example is for an array of rectangular waveguides arranged in a triangular lattice, both with and without conducting "irises" that block the left and right 1/4 of each waveguide aperture. The discrete points shown in Figure 4 are HFEM calculations and the solid lines are multimode calculations by Lee & Jones [17]. The two methods agree on the approximate location of the scan blindness without irises near 33°, and show that the effect of adding the irises is to suppress the scan blindness, at the expense of approximately 2 dB of broadside gain. For this test case, the finite element

271
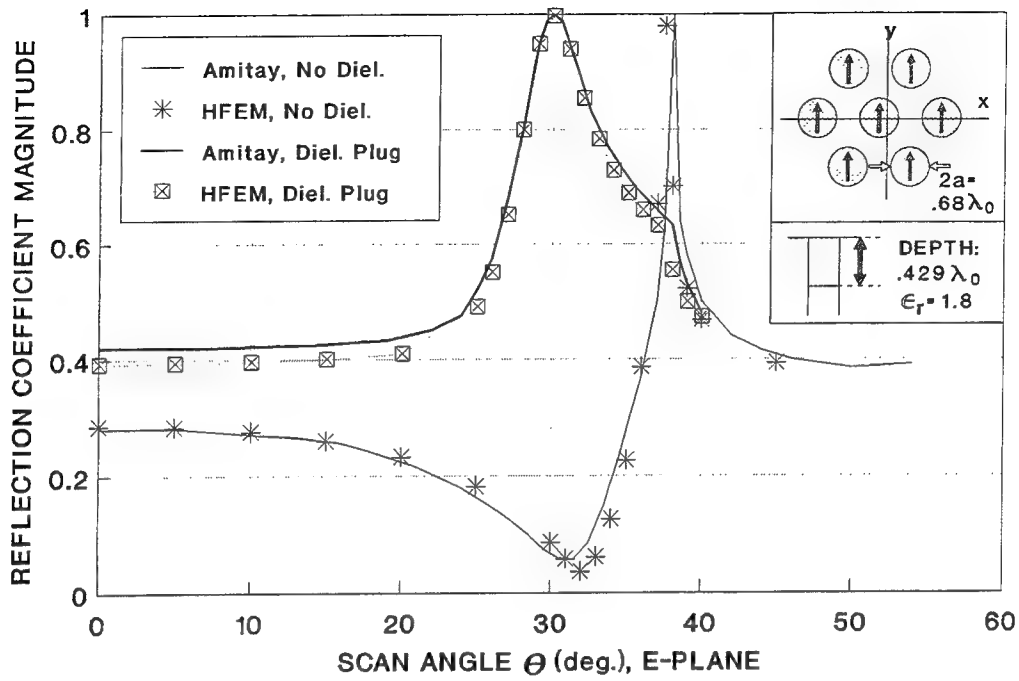
Figure 3. Reflection Coefficient vs. Scan Angle, Circular Waveguide Array: radius a$=.34\lambda_0$, $d_x=.714\lambda_0$, $d_y=.618\lambda_0$, $\gamma=60°$, plug depth h$=.429\lambda_0$, $\epsilon_r=1.8$ [2].
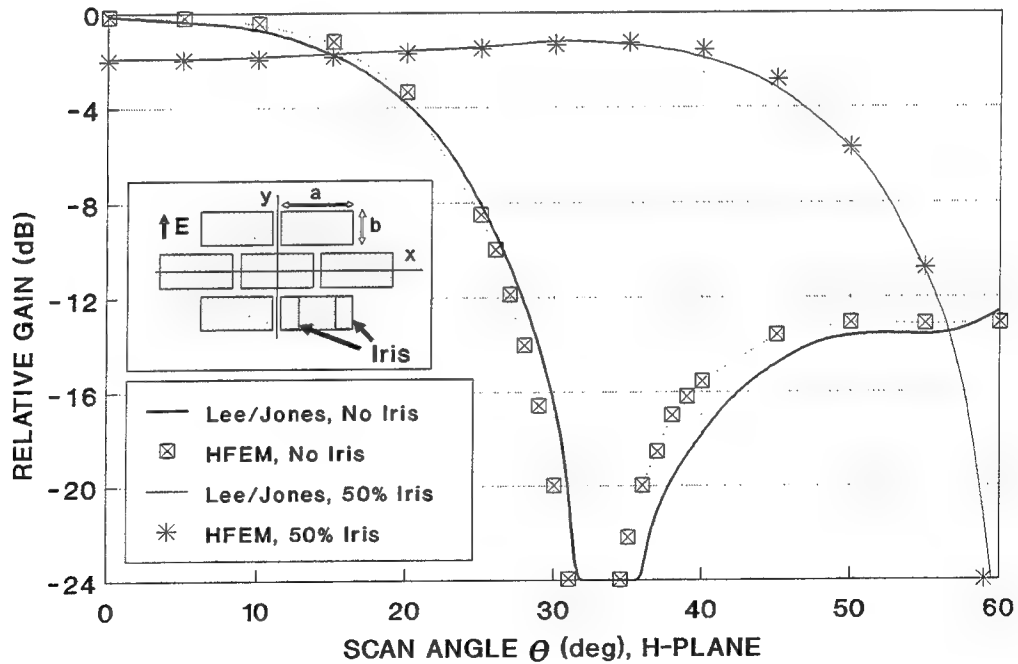
Figure 4. Active Element Gain vs. Scan Angle, Rectangular Waveguides: width a$=.905\lambda_0$, height b$=.4\lambda_0$, $d_x=1.008\lambda_0$, $d_y=.504\lambda_0$, $\gamma=45°$ [17].

272

model was simply a short section of rectangular waveguide.

A similar example case that illustrates some of the new capabilities using the finite element method is shown in Figure 5. The radiating element is again an open-ended rectangular waveguide, except it includes a transition from coaxial waveguide. This end-wall waveguide launcher was discussed by Tang & Wong [18] as a feed for a waveguide array element, although they placed it far enough behind the waveguide apertures that any evanescent modes it excited would not be able to radiate. In this example, however, the launcher is quite close to the aperture. Figure 6 shows the HFEM calculations, and includes the semi-infinite waveguide feed for comparison. The launcher's proximity to the aperture is evidently preventing the scan blindness from occurring, at the expense of less than 1 dB in broadside gain.
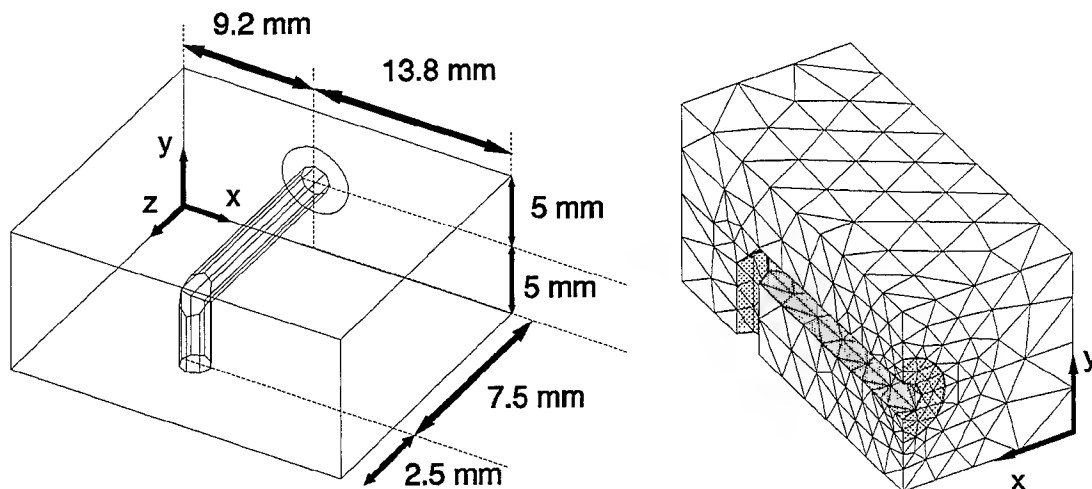
Figure 5. Rectangular Waveguide Radiator with Coaxial End-wall Launcher [18]:
(a) Geometry; (b) Cutaway of Tetrahedron Mesh
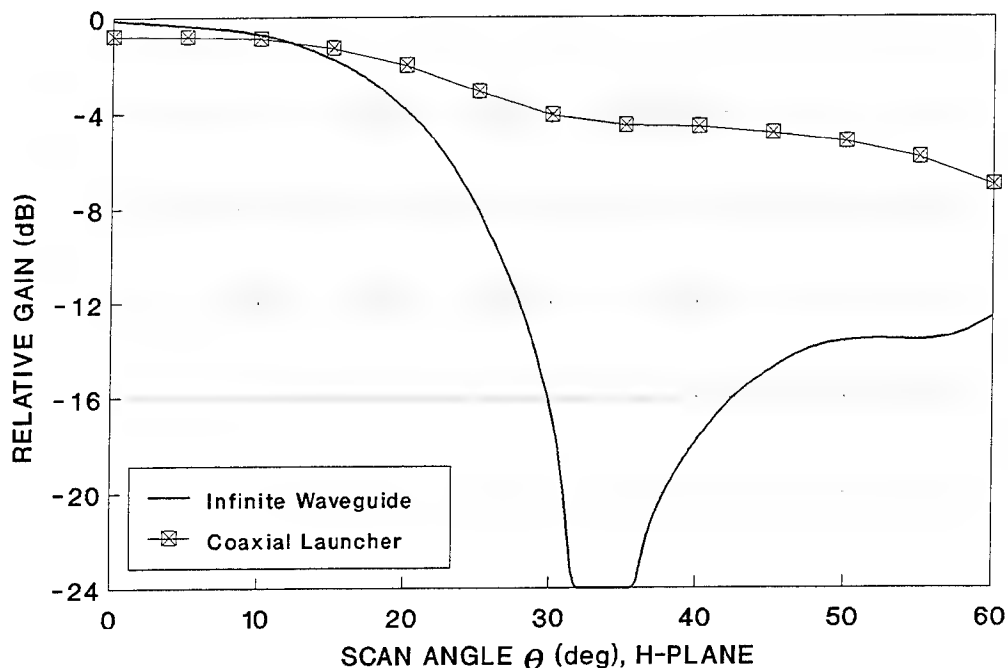
Figure 6. Element Gain vs. Scan Angle for Rectangular Waveguide Radiator
with Coaxial Launcher near Aperture

3.2 Pyramidal Horn Array

Figure 7 shows the geometry of a pyramidal horn used in an array constructed

by Amitay & Gans [19]. Also shown is the finite element mesh with one quadrant

removed. Note that the mesh does not extend into the waveguide, but terminates at the

junction with the horn flare. This case required more Floquet modes, 61 and 31 in the

k_x and k_y, respectively, because of the number of modes required is proportional to the

size of the aperture in wavelengths. Figure 8 compares the measured element gain with

calculations using HFEM.

In the preceding cases, the radiators only couple through apertures in a common

ground plane, so the matrix transformation (24) was not required. In contrast, the

examples that follow are cases in which the radiator structures project above the common
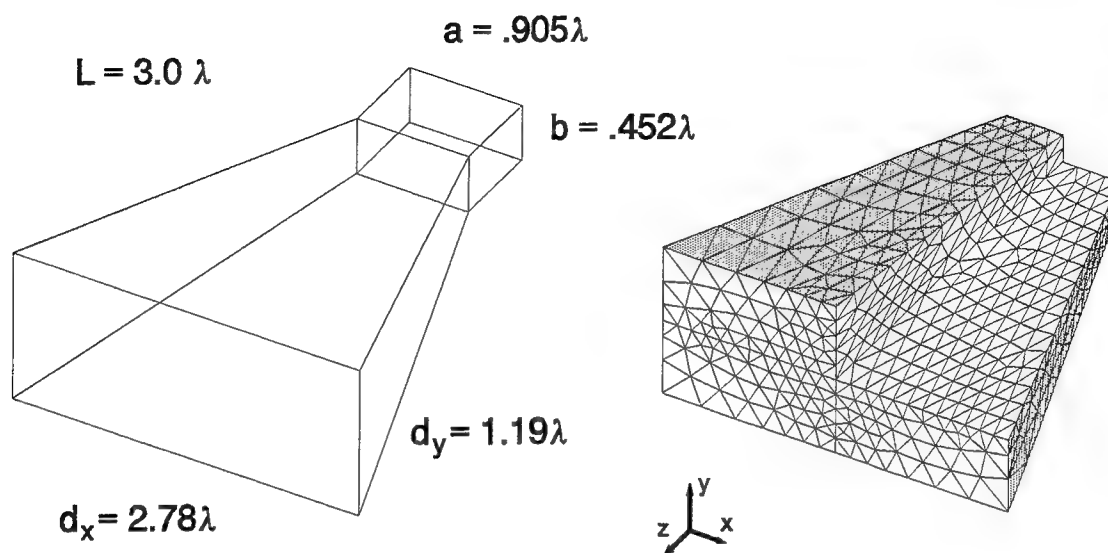
Figure 7. Pyramidal Horn Element Geometry and Finite Element Model
with One Quadrant Removed
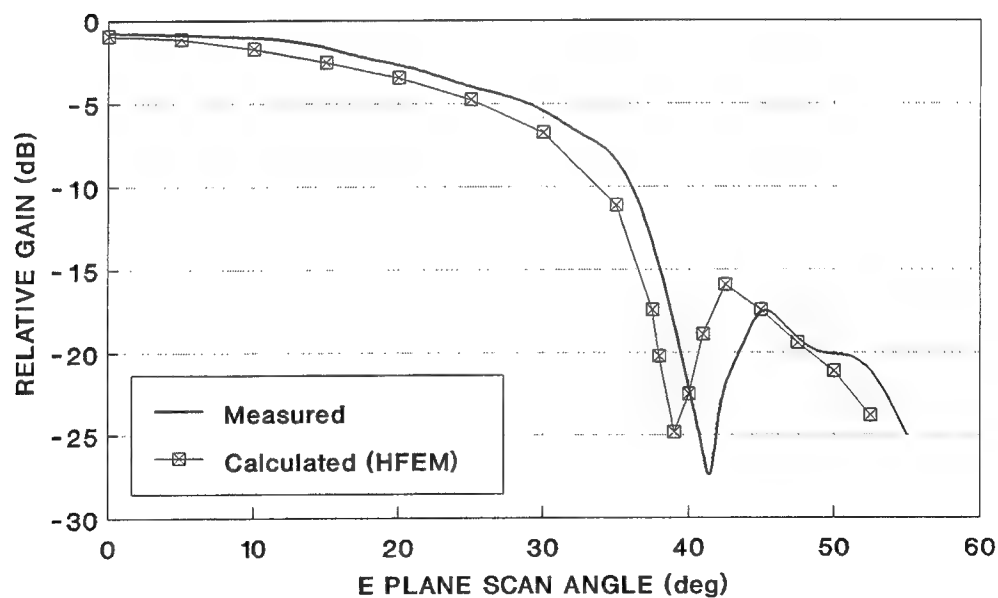
Figure 8. Element Gain Calculation vs. Measurement [19]
for Pyramidal Horn Array

ground plane, so there is also coupling across unit cell side walls.

3.3 Microstrip Arrays

Two cases of rectangular microstrip patch arrays on high-permittivity substrates ($\epsilon_r=12.8$, representing Gallium Arsenide) were modeled by Pozar and Schaubert [20] with the method of moments to demonstrate surface wave effects. The two substrate thicknesses were $h=.02\lambda_0$ (thin case) and $h=.06\lambda_0$ (thick case). In the finite element model, the patch feed pin is an extension of the coaxial feed's center conductor. The finite element model consists of a rectangular section of the dielectric slab, and the patch coincides with the boundary Γ_R.

Figure 9 is a comparison of the MoM calculations and the HFEM calculations. It shows the *normalized* active reflection coefficient, emulating the effect of a matching network that adjusts for zero reflection coefficient at broadside scan:

$$R_{norm} = \left| \frac{Z_{in}(\theta,\phi) - Z_{in}(0,0)}{Z_{in}(\theta,\phi) + Z_{in}(0,0)} \right| \tag{31}$$

The patch dimensions are $.15\lambda_0$ x $.131\lambda_0$ and $.15\lambda_0$ x $.098\lambda_0$ for the thin and thick cases, respectively. Both MoM and HFEM are predicting the existence and approximate angular location of blindnesses due to surface waves in the dielectric slabs. The slight disagreement for the thick case is due to a difference in the feed model: while the MoM used an idealized probe, the HFEM used a coaxial aperture (inner and outer radii of $.0075\lambda_0$ and $.0175\lambda_0$, respectively). Aberle and Pozar [21] showed that when an accurate coaxial feed model is included, the MoM calculations also give a reflection coefficient that is slightly below 1.0 at the blindness angle.

276
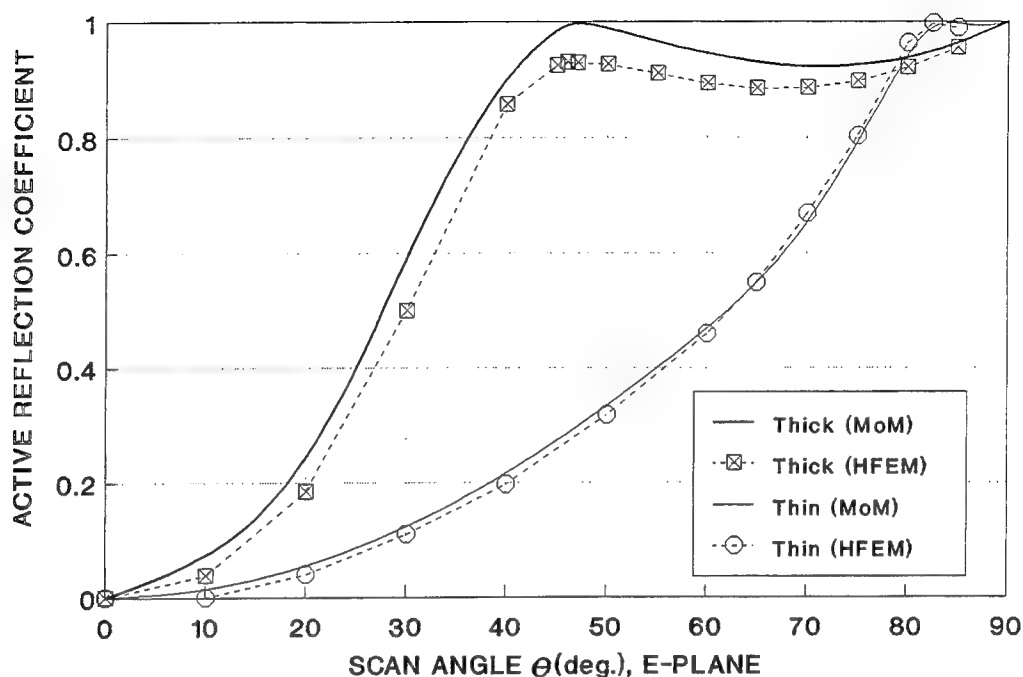
Figure 9. Comparison of HFEM and MoM [20] Calculations for Microstrip Patch Arrays on $\epsilon_r = 12.8$ Substrate, Thickness $= .02\lambda_0$ and $.06\lambda_0$, $d_x = d_y = .5\lambda_0$

3.4 Clad Monopole Array

The monopole radiator is simply an extension of a coaxial waveguide. Figure 10 shows a cutaway of the finite element model used for this example. (Since the array lattice is skewed, one side of the unit cell is not parallel with a coordinate axis.) The model includes an extension of the coaxial line's dielectric insulator, forming a cladding, or sleeve, around the monopole. The purpose of this dielectric cladding around the monopole is to enhance its bandwidth. Calculations using this element have shown an increase of over 50% in bandwidth over the bare monopole with a Teflon sleeve ($\epsilon_r = 2.1$). Figure 11 shows the HFEM calculations for active reflection coefficient vs. scan angle compared to measurements made on a 9 x 11 element finite array. The

277
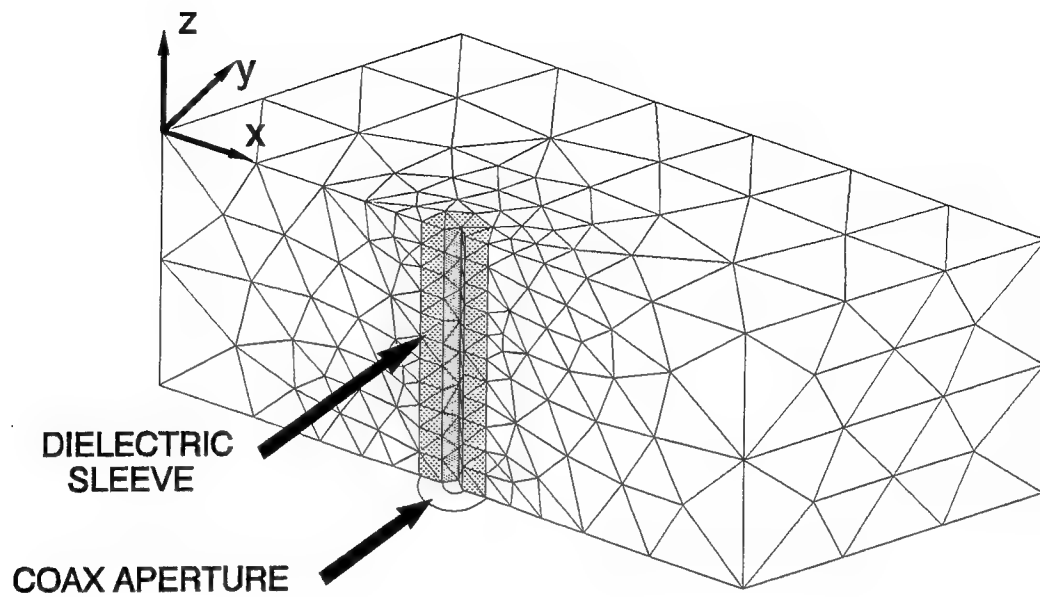
Figure 10. Cutaway of Finite Element Model for Clad Monopole Radiator
$d_x = .576\lambda_0$, $d_y = .496\lambda_0$, $\gamma = 60°$, monopole length $= .235\lambda_0$, sleeve length $= .189\lambda_0$
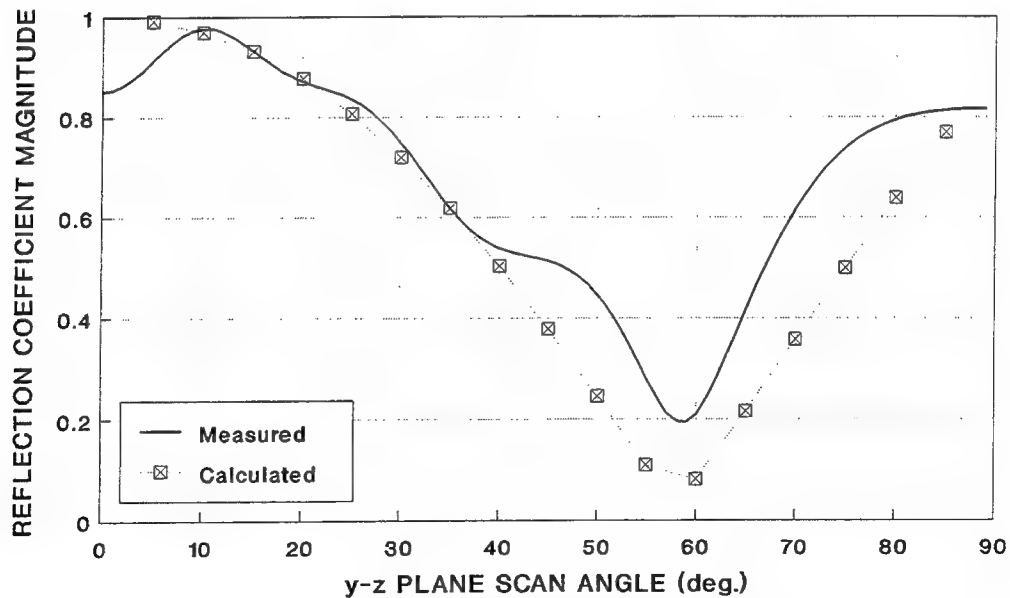
Figure 11. Calculated and Measured Reflection Coefficient Magnitude
vs. Scan Angle for Clad Monopole Array

monopole elements were ordinary SMA coaxial connectors whose inner and outer radii correspond to $.010\lambda_0$ and $.033\lambda_0$, respectively. The active reflection coefficient was calculated from the measured mutual coupling using

$$R_a(\theta,\phi) = \sum_{m=-\infty}^{\infty} \sum_{n=-\infty}^{\infty} C_{mn}\, e^{-j\psi_x(md_x+nd_y\cot\gamma)}\, e^{-j\psi_y nd_y} \qquad (32)$$

where C_{mn} is the measured S_{21} between a reference element at the center of the array and the element in row m and column n. The finite array measurements only give an approximation to R. For that reason, the calculations and measurements are not expected to agree completely, especially near endfire ($\theta=90°$).

3.5 Printed Flared Notch Array

The final test case, a printed flared notch, is representative of the initial motivation for this work. The design shown in Figure 2 uses a novel coplanar waveguide (CPW)-slotline transition by Ho & Hart [22] to restrict metallization to one side of the substrate only. The overall element height is 38.1 mm, and the length of the flared section is 33.3 mm. The coaxial waveguide feeds a short section of coplanar waveguide that transitions to slotline by terminating one of the parallel slots in a broadband open. The slotline then leads into an exponential flare [23]. Note that since the substrate has a much higher permittivity ($\epsilon_r=6$) than the surrounding air, the mesh is much finer in the substrate, and gradually relaxes going out towards the +x and -x unit cell side walls. The mesh is made even finer near the coax-CPW and CPW-slotline transitions to accurately capture the details of the geometry.

The HFEM predictions for active reflection coefficient vs. scan angle are shown

279
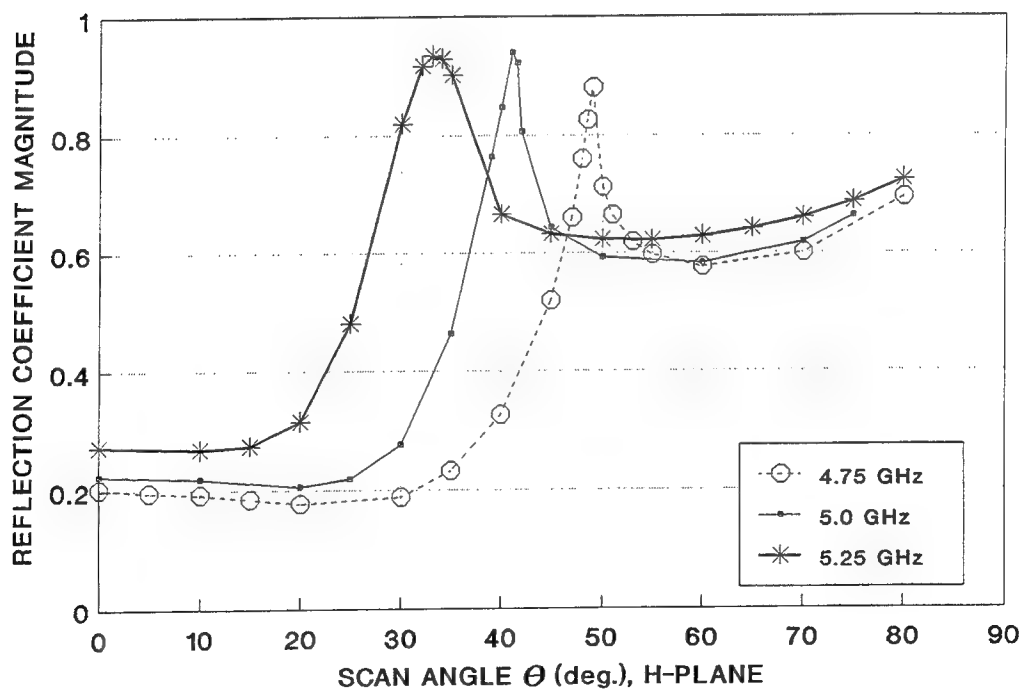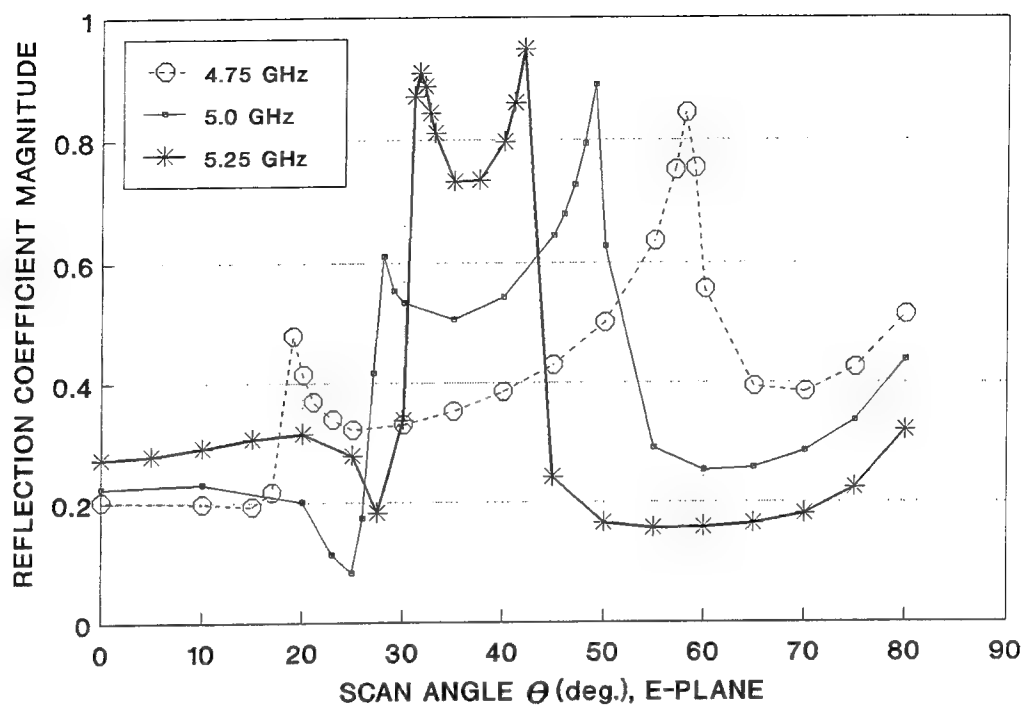
Figure 12. HFEM Predictions of Active Reflection Coefficient vs. Scan Angle for Flared Notch Array: $d_x = 36$ mm, $d_y = 34$mm, $\gamma = 90°$

in Figure 12. The radiator is well-matched to the array environment, with reflection coefficient less than .3 from 3.0 to 5.0 GHz. Scan blindnesses occur at the grating lobe angle in both E and H planes when the frequency exceeds that at which the inter-element spacing is more than one half wavelength (4.4 GHz in the E plane; 4.2 GHz in the H plane). There is an additional E plane blindness that moves *outward* in angle with increasing frequency. Schaubert and Aas have also observed this unusual behavior in moment method calculations, and have shown that it is due to the excitation of a waveguide mode on the corrugated structure [24].

3.6 Computer Storage and Execution Time

The computer storage required is dominated by the matrix, whose size depends on the total number of nonzero matrix elements. The matrix is stored in three parts: the sparse matrix due to the finite element interactions; and two dense matrices due to the waveguide and periodic radiation conditions. For the sparse portion, it is only necessary to store the nonzero matrix entries since the solution is by iterative methods. This computer code performs all calculations in single precision. The total storage in bytes is approximately $10(12N)+(8N_W)^2+(8N_R)^2$, where N is the total number of edges, N_W is the number in the waveguide aperture, and N_R is the number in the radiation boundary. There are approximately 10 nonzero entries in each matrix row. A single complex number is 8 bytes, but for the sparse matrix entries there are two integer arrays holding the row and column indices. The maximum storage required by any of the foregoing test cases was 45 Mbytes, which was the microstrip patch.

The execution time is usually dominated by the matrix solve. With both the

281

conjugate gradient (CGM) and biconjugate gradient methods (BCGM), the time consuming part of the computation is the matrix-vector multiply. Each iteration requires two of these for a general non-symmetric matrix. Hence, the total number of matrix elements is the most important factor in execution time.

The execution times (on a SparcStation^{TM}-10) for the more challenging test cases are presented in Table I, broken down by fill time (computing the matrix entries) and solve time for CGM and BCGM. Each new frequency or scan angle requires a new calculation of S^R and a complete matrix solve. Both solution methods converged for all test cases, but BCGM had an advantage of between 1.5 and 10 in speed. The solve time was the time required to reduce the residual norm to .001 of its initial value (with the solution vector set to zero). The microstrip patch was by far the most time consuming

Table I. Matrix Size and Execution Times for Example Problems
(Time is CPU minutes on SparcStation^{TM}-10)

	Rect. WG & Coaxial Launcher	Pyramidal Horn	Clad Monopole	Thick Patch	Printed Flared Notch
Number Edges (N)	3669	8466	3487	4862	5271
Edges in Γ_W	72	72	40	12	15
Edges in Γ_R	75	442	187	841	181
Matrix Fill Time	1.7	177	2.7	18.5	4.0
CGM Solve Time	37	55	89	118	71
CGM Iterations	11524	3289	20000	2906	12227
BCGM Solve Time	7.0	33	11	86	20.5
BCGM Iterations	2054	1878	2340	2107	3512

case in terms of solution time. This is due to the fact that there are so many edges in Γ_R, so the dense portion of the matrix is quite large (841^2 words). The fill time was greatest for the pyramidal horn because of the large number of Floquet modes. These statistics illustrate that the code can produce results for practical problems, even with complicated geometries, within a reasonable time on a moderately-powerful workstation.

4. CONCLUSIONS

The finite element method was applied to the analysis of phased array radiators. This required developing boundary conditions for waveguide feeds, a periodic radiation condition, and side wall periodicity conditions. The first of these provides an accurate means of modeling rectangular, circular and coaxial waveguide feeds. The second and third allow accurate modeling of radiation from elements in large (assumed infinite) arrays, including scan blindness effects. The validation and demonstration cases showed that the techniques are valid for radiators with complex feed structures, irregular conductor configurations and inhomogeneous dielectrics. Finally, the fact that the same computer code was used for all of these cases demonstrates the unprecedented versatility obtainable using the hybrid finite element method.

5. ACKNOWLEDGEMENTS

6. REFERENCES

[1] Stark, L. "Microwave Theory of Phased-Array Antennas - A Review," *Proc. IEEE*, **62**, pp. 1661-1701, Dec. 1974.

[2] Amitay, N., V. Galindo and C. Wu, *Theory and Analysis of Phased Array Antennas*, New York: Wiley, 1972.

[3] Jin, J-M. and V. V. Liepa, "Application of the Hybrid Finite Element Method to Electromagnetic Scattering from Coated Cylinders," *IEEE Trans. Antennas Propagat.*, **AP-36**, pp. 50-54, Jan. 1988.

[4] D'Angelo, J.D. and I.D. Mayergoyz, "Finite Element Methods for the Solution of RF Radiation and Scattering Problems," *Electromagnetics*, **10**, pp. 177-199, 1990.

[5] Boyse, W.E. and A.A. Seidl, "A Hybrid Finite Element and Moment Method for Electromagnetic Scattering from Inhomogeneous Objects," *7th Annual Review of Progress in Applied Computational Electromagnetics*, Monterey CA: Naval Postgraduate School, pp. 160-169, Mar. 1991.

[6] Barton, M.L. and Z.J. Cendes, "New Vector Finite Elements for Three-Dimensional Magnetic Field Computation," *J. Appl. Phys.*, **61**, pp. 3919-3921, Apr. 1987.

[7] Silvester, P.P. and R.L. Ferrari, *Finite Elements for Electrical Engineers*, 2nd. ed., Cambridge Univ. Press, 1990.

[8] Mur, G., "Finite-Element Modeling of Three-Dimensional Electromagnetic Fields in Inhomogeneous Media, *Radio Science*, **26**, pp. 275-280, 1991.

[9] Harrington, R.F. and J.R. Mautz, "A Generalized Network Formulation for Aperture Problems," *IEEE Trans. Antennas Propagat.*, **AP-24**, pp. 870-873, Nov. 1976.

[10] Marcuvitz, N., *Waveguide Handbook*, New York: McGraw-Hill, 1951.

[11] McGrath, D.T., "Hybrid Finite Element/Waveguide Mode Analysis of Passive RF Devices," Griffiss AFB, NY: Rome Laboratory, RL-TR-93-130, June, 1993. **ADA273074**

[12] McGrath, D.T., "Phased Array Antenna Analysis Using Hybrid Finite Element Methods," Ph.D. Dissertation, Wright-Patterson AFB, OH: Air Force Institute of Technology, June, 1993.

[13] Gaskill, J.D., *Linear Systems, Fourier Transforms, and Optics*, New York: J. Wiley & Sons, 1978.

[14] Gedney, S.D., J.F. Lee and R. Mittra, "A Combined FEM/MoM Approach to Analyze the Plane Wave Diffraction by Arbitrary Gratings," *IEEE Trans. Microwave Theory Tech.*, **MTT-40**, pp. 363-370, Feb. 1992.

[15] Sarkar, T.K., "On the Application of the Generalized Biconjugate Gradient Method," *J. Electromagnetic Waves Appl.*, **1**, pp. 223-243, 1987.

[16]. McGrath, D.T., and V.P. Pyati, "Hybrid Finite Element/Waveguide Mode Analysis for Infinite Phased Arrays of Cavity Radiators," *1993 IEEE AP-S Symposium Digest*, pp. 1242-1245, June, 1993.

[17] Lee, S-W and W. Jones, "On the Suppression of Radiation Nulls and Broadband Impedance Matching of Rectangular Waveguide Phased Arrays," *IEEE Trans. Antennas Propagat.*, **AP-19**, pp. 41-51, Jan. 1971.

[18] Tang, R. and N.S. Wong, "Multimode Phased Array Element for Wide Scan Angle Impedance Matching, *Proc. IEEE*, **56**, pp. 1951-1959, Nov. 1968.

[19] Amitay, N. and M.J. Gans, "Design of Rectangular Horn Arrays with Oversized Aperture Elements," *IEEE Trans. Antennas Propagat.*, **AP-29**, pp. 871-884, Nov. 1981.

[20] Pozar, D.M. and D.H. Schaubert, "Analysis of an Infinite Array of Rectangular Microstrip Patches with Idealized Probe Feeds," *IEEE Trans. Antennas Propagat.*, **AP-32**, pp. 1101-1107, Oct. 1984.

[21] Aberle, J.T. and D.M. Pozar, "Analysis of Infinite Arrays of One- and Two-probe-fed Circular Patches," *IEEE Trans. Antennas Propagat.*, pp. 421-432, Apr. 1990.

[22] Ho, T.Q. and S.M. Hart, "A Broad-Band Coplanar Waveguide to Slotline Transition," *IEEE Microwave and Guided Wave Letters*, **2**, pp. 415-416, Oct. 1992.

[23] Choung, Y.H. and C.C. Chen, "44 GHz Slotline Phased Array Antenna," *1989 IEEE AP-S Symposium Digest*, pp. 1730-1733, June, 1989.

[24] Schaubert, D.H. and J.A. Aas, "An Explanation of Some E-Plane Scan Blindnesses in Single-Polarized Tapered Slot Arrays," *1993 IEEE AP-S Symposium Digest*, pp. 1612-1615, June, 1993.

EFFICIENT MODELING OF MICROSTRIP ANTENNAS USING THE FINITE DIFFERENCE TIME DOMAIN METHOD

Siva Chebolu, John Svigelj, and Raj Mittra
Electromagnetic Communication Laboratory,
Department of Electrical and Computer Engineering,
University of Illinois, Urbana, IL 61801-2991.
Phone: (217) 333-1200 FAX: (217) 333-8986

Abstract

While complex microstrip antenna configurations can be conveniently analyzed using the Finite Difference Time Domain (FDTD) method, accurate modeling of the fine features of these antennas with this approach typically requires a large mesh, and an exorbitant amount of computational time. In this paper, we describe some efficient FDTD techniques for obviating these difficulties that arise when modeling such geometries. We show that significant savings in the computation time can be achieved by applying extrapolation techniques to these resonant structures. In addition, an appropriate choice of the excitation allows us to obtain the radiation and impedance characteristics of the antennas over the desired range of frequencies with a single simulation. Furthermore, knowledge of the current distribution on the patch provides physical insight into the behavior of these antennas.

Numerical results illustrating the application of the FDTD approach to a two-layer stacked patch antenna are presented in the paper.

1.   Introduction

Over the last decade, microstrip antennas have been a topic of extensive research and development, and they are increasingly being used in a wide variety of applications ranging from satellite communication to biomedical systems, typically in the frequency range of 500 MHz to 30 GHz. Some of the attractive features of a microstrip antenna are: low profile, lightweight nature, conformability to curved surfaces, ease of fabrication using photolithographic techniques, low cost, and ease of integration with feeding networks and associated

circuitry. However, these antennas do suffer from relatively low bandwidth, low polarization purity, spurious feed radiation, low power handling capabilities and tolerance problems. For a discussion of these attributes of microstrip antennas and good reviews of current microstrip antenna technology the reader is referred to [1] and [2].

Microstrip antenna systems can be broadly classified into two categories:

(i)   Large microstrip arrays radiating high levels of power, as in satellite transmitting antennas.

(ii)   Receiving or transmitting antennas, for low power applications in biomedical probing and cellular or GPS communications. These structures are typically characterized by a single microstrip antenna or by an array with relatively few elements.

In this work, we will focus our attention to the problem of modeling microstrip antennas that belong to the second category.

In recent years, a wide variety of techniques have been developed to analyze microstrip antennas and arrays. Simple models, such as the cavity model [3], and the transmission line model [4], can be used to predict the antenna characteristics and to gain an intuitive understanding of the radiation mechanism of the antenna. However, the accuracy of these approximate models is fairly limited and they are only suitable for analyzing simple, regularly-shaped microstrip antennas on thin substrates. An alternative to using the cavity model is to employ full-wave methods that allow us to handle arbitrarily-shaped antennas with either a coaxial or a microstrip feed structure. One such technique is the Method of Moments (MoM), which has been extensively used in microstrip

antenna analysis, both in the spectral [5] and spatial [6] domains. Although this method is well-suited for handling planar microstrip structures mounted on large ground planes, and is capable of modeling a variety of feed structures, the MoM formulation requires extensive preprocessing to construct the requisite Green's functions for the given geometry. In addition, an improper choice of the basis functions in MoM can result in an ill-conditioned matrix, whose solution may pose convergence problems. This is especially true when the antenna geometry is complex, and the ground plane has a finite size.

Another full-wave modeling tool, which is being increasingly used for a wide range of applications, e.g., modeling of electronic circuits, electromagnetic interference, electromagnetic compatibility, and radar scattering, is the Finite Difference Time Domain (FDTD) method [7-9]. Recently, it has also been applied to the analysis of a class of microstrip structures [10-15], and it has been shown that this method can not only conveniently handle a variety of microstrip antennas, but can also model the matching networks and associated circuitry feeding the antenna with relative ease. Numerically rigorous treatment of antennas with several dielectric layers, and of various feed structures such as coaxial, microstrip and aperture excitations, can be carried out using this approach. The method can also be used to analyze active antennas containing non-linear elements, and, since the method solves the problem directly in the time domain, it enables us to determine the stable mode of operation of such antennas [14]. Another attractive feature of the FDTD method is that it requires very little preprocessing and hence a generic user-friendly code can be developed for analyzing a wide variety of structures.

However, one major limitation of the conventional FDTD algorithm [16] is that modeling of the fine features of the microstrip antenna geometry places a heavy burden on the CPU time and memory requirements. This is because computation over many time steps is required to attain convergence owing to the resonant nature of these structures and the fine resolution of the mesh.

In this paper, we summarize several techniques that enhance the computational efficiency of the FDTD algorithm and illustrate the application of these techniques via an example of a two-layer stacked patch antenna.

2.   FDTD modeling of microstrip antennas

Several issues, discussed below, must be examined in the process of FDTD modeling of microstrip antennas. These are:

- *Spatial discretization*

  To accurately model the fine features of the microstrip antenna, such as the thickness of the substrate and the feed structure, it is often necessary to use spatial discretizations that are on the order of $\lambda_0/30$ - $\lambda_0/100$, much smaller than the nominal value of $\lambda_0/10$, where λ_0 is the wavelength in the medium at the highest frequency of interest. The Courant stability condition [7], applied to the fine mesh, forces us to use a time discretization which is smaller by a factor of 3 to 10 as compared to the conventional time step. In summary, modeling of antennas with fine features using a uniform orthogonal grid requires a large computational mesh and a correspondingly large number of time iteration steps, and the process becomes both memory and CPU-time intensive.

To obtain an accurate field solution in an arbitrary structure by using the FDTD algorithm, it becomes necessary to choose a suitable grid discretization such that it accurately models both the geometry and the field variations in the structure, and one approach to doing this is to use a general, finite-element type of non-orthogonal grid [17]. However, this approach is not very robust as it is not suitable for handling different structures with a minimum amount of preprocessing, which is one of the principal attractive features of the FDTD method. An alternative strategy is to use the curvilinear FDTD approach, which employs a structured non-orthogonal grid, and is well-suited for modeling patch antennas with curved surfaces [18]. However, for the class of antennas analyzed in this work, the curvilinear method turns out to be prohibitively expensive in terms of memory requirements and CPU time. This is because it is necessary in the aforementioned approach to store the coordinate transformation metrics, and because the time required to convert the covariant and contravariant field components back and forth is computationally expensive. Fortunately, for our application, the only curved geometries are the circular coaxial feeds; hence, simply approximating these lines with square geometries that have the same impedance, yields good results and obviates the need to use the curvilinear FDTD. However, we can still improve the modeling accuracy of the problem at hand by using a non-uniform but orthogonal grid, which offers a good compromise between the uniform and non-orthogonal discretizations.

There are several advantages to using the FDTD on a non-uniform orthogonal grid. First, the orthogonality preserves the simplicity of the form of the FDTD equations without significantly burdening the memory requirements or compromising the speed of computation of the uniform FDTD. Second, the flexibility of the non-uniform mesh allows us to model the geometry more accurately than is possible in the uniform approach which requires the dimensions of a structure to be integral multiples of the chosen discretizations in the x, y, and z directions. Third, the non-uniform orthogonal grid provides the flexibility to model the fine features of a geometry with a denser mesh in the region of interest, where the field variation is rapid, and a coarser mesh elsewhere. Variability of the mesh resolution also enables us to move the boundaries of the computational domain farther away from the center of the problem, without an undue increase in the number of cells. This can be quite important when modeling large, complicated structures, because the presence of spurious reflections from imperfect absorbing boundary conditions (ABCs) can contaminate the FDTD solution when these boundaries are emplaced too close to the radiating or scattering structure.

• *Source Excitation*

Microstrip antennas are inherently high-Q structures, and are typically designed to operate over a narrow band of frequencies. To obtain the frequency information within this narrow range, it is usual in the FDTD analysis to excite the antenna with a sinusoidal signal, operating at the center frequency of the band, and modulate it by an appropriate Gaussian

291

pulse [10]. This type of temporal excitation not only excites just the desirable modes on the antenna, but also allows the fields to settle down quickly and achieve the steady state of the dominant, resonant modal distributions. The situation can be further improved by replacing the Gaussian pulse with a windowed one, say with a Blackman-Harris window, whose spectral response has very low side lobes.

Typically, microstrip antennas are excited using a coaxial probe or a microstrip feeding structure which could be coupled to the patch via an aperture. In either case, it is numerically efficient to excite these transmitting lines with a quasi-static solution obtained from the solution of the Laplace's equation. With this form of spatial excitation, the fields settle down relatively quickly to their modal distributions and, hence, it is possible to use a shorter length of the transmission line for modeling the feed structure.

- *Extrapolation techniques*

Owing to the resonant nature of the microstrip antennas, the transient response of the fields typically take a long time to converge. However, this ringing feature of the temporal response is also ideally suited for extrapolation because it is relatively straightforward to extract the resonant frequencies from the time domain signature over a short duration, and subsequently use this information to extrapolate this signature. Various procedures, such as the generalized pencil-of-function (GPOF) method [19], or the auto-regressive (AR) models [20], can be employed for this purpose.

- *Input Impedance*

  The input impedance is computed by monitoring the voltage V(t) and the current I(t) at suitable locations along the feed. Fourier transformation of the time domain data, followed by a calculation of the voltage to current ratio V(f)/I(f) in the frequency domain, yields the variation of the input impedance Z_{in} over the desired frequency range. These results are then translated to a suitable reference plane, determined by the measurement settings, and compared with the experimental data.

- *Radiation Pattern*

  The radiation pattern over the desired frequency range can be obtained with a single simulation using the time domain near- to far-zone transformation [21]. For a single microstrip antenna mounted on an infinite ground plane, we can use the temporal current distribution on the metal patch to calculate the radiation pattern [10]. However, for more complex geometries involving two or more stacked patches, it is convenient to monitor the fields on a suitable equivalent surface enclosing the radiating structure and then process this information to obtain the radiation pattern.

- *Distributed Computing*

  Memory limitations on present day workstations typically allow us to handle only simple structures. This difficulty can be partially overcome by using a domain decomposition procedure wherein the original problem is subdivided into several appropriate sub-problems. For example, in a probe-fed microstrip antenna, the coaxial connector occupies a space

which is much smaller than that of the microstrip patch. Hence, we can save significant amount of computational resources by dividing the problem into two parts consisting of the microstrip patch and the coaxial feed [12]. These two problems can then be run separately with the field values at the interface being communicated between the two regions via a message passing scheme. Finally, it is possible to handle large complex structures by using distributed computing techniques [22], that are especially suited for microstrip geometries which can be partitioned relatively easily.

- *Other issues*

For accurate modeling of the characteristics of the antenna, we need to account for the loss tangent (tan δ) of the dielectric, as it can significantly alter the magnitude of the input impedance [23]. This feature can be easily incorporated into the FDTD algorithm with only a slight modification of the code by introducing an equivalent conductive loss, given by $\sigma_{eff} = \omega \, \varepsilon \tan \delta$, in the dielectric. Note that an appropriate value of ω, which lies within the given frequency range, must be used since the FDTD approach is not too-well suited for handling dispersive media.

3.   The two-layer stacked patch antenna

Recently, non-conformal microstrip antennas involving multiple layers of stacked patches have found applications in GPS communications [24]. In this section we consider the FDTD modeling of a two-layer stacked patch geometry, shown in Fig. 1, which is a circularly-polarized (CP) antenna used in mobile

communications for realizing a near-omnidirectional coverage [24]. We are interested in obtaining the characteristics of this antenna in the frequency range of 1-2 GHz, centered about its first two resonances.

The geometry is first discretized using the non-uniform rectangular brick elements as shown in Fig. 2(a). As an illustration of the non-uniform orthogonal mesh, a top view of the FDTD grid for the bottom patch is shown in Fig. 2(b). The smallest discretization is used in modeling the inner coaxial connectors and the grid is gradually expanded. The growth factor of the mesh is kept lower than 1.3 to prevent artificial discontinuities due to abrupt changes in cell size. The mesh used to describe the two-layer patch antenna structure had 106 x 107 x 126 cells, for the computational domain of dimensions 244 mm x 244 mm x 225 mm. The smallest cell was 0.75 mm x 0.75 mm x 0.75 mm, resulting in a time step of 1.25 ps, and the largest cell was 3.2 mm x 3.2 mm x 3.2 mm. Note that at the highest frequency of interest, viz., 2 GHz, the largest cell discretization is on the order of $\lambda_0/47$. Using a uniform mesh with a cell size of 1 mm x 1 mm x 1.6 mm to model the same computational domain would require a mesh with 244 x 244 x 140 cells. Compared to the uniform discretization for this case, we see that the use of the non-uniform grid results in a memory savings of 83%.

For the first two resonant modes of the bottom patch, the electric fields inside the dielectric are zero are the center of the patch. Hence, although the feed for the top patch passes through the center of the bottom patch it does not significantly alter the field distribution on the bottom patch. The top patch is fed about one-third way along its diagonal to achieve CP with a good impedance match. Therefore, we need to bend the coaxial cable that feeds the top patch as

shown in Fig. 1. This bend is modeled in FDTD using two 90 degree bends as shown in Fig. 2 (a).

The coaxial feeds are terminated using the first-order Mur ABC [25], while the remaining unbounded regions of the computational mesh are truncated using the second-order Mur ABCs. The feeds are excited with a spatial distribution corresponding to the quasi-static Laplace's solution for the coaxial line. The temporal form of the excitation is shown in Fig. 3(a) and its frequency spectrum is shown in Fig. 3(b). The two patches are fed in phase using a two-way power divider. This effect is simulated in the FDTD modeling by exciting both the patches with the same form of input pulse at locations along the feeds that are equidistant from the patches.

Figs. 4 (a) and 4 (b) show the time-domain signatures of the voltage and the current which are monitored at the intersection of the feed and the patch. Then, the GPOF method is used on the data from 4000-6000 time steps to extrapolate the response to 25000 time steps. As seen in Fig. 5, there is excellent agreement between the actual data and the extrapolated data. Post processing the extrapolated data yields the input impedance. Time-domain data over about 25000 time steps are required to ensure accurate frequency domain results. This task was accomplished through application of the GPOF extrapolation technique which employs data only over 6000 time steps, and saves about 76% in terms of CPU time.

The input impedance of the bottom patch is shown in Fig. 6(a). The two resonances in the impedance curve at $f_1 = 1461$ GHz and $f_2 = 1491$ GHz correspond to the first two dominant modes of the nearly square antenna. These

frequencies differ from the experimentally measured values by less than 4%. Fig. 6(b) shows an alternative way of displaying the impedance information. The presence of the cusp in the impedance plot indicates the circularly-polarized operation of the antenna.

Fig. 7 shows the field distributions on the bottom patch, operating in isolation, after the response has stabilized to its steady state value. The plot of the electric field, E_z, exhibits both the TM_{01} and TM_{10} modal distributions. We note from the plots for the current distribution that there is a charge accumulation along the edges of the patch.

4. Conclusions

FDTD is a simple, accurate, and versatile technique for modeling various types of microstrip structures. It is well-suited for handling the complex geometries, and for conveniently generating, with a single simulation, information on the various characteristics of the antennas, such as the input impedance and the radiation pattern over the desired frequency range. Also, the ability to visualize the field and current distributions on the patches is helpful in developing an understanding of the mechanism of radiation from the antenna.

We have described several techniques that can help speed up the computational efficiency of the FDTD algorithm modeling complex microstrip antennas. These techniques include choosing a proper type of feed excitation; employing appropriate absorbing boundary conditions; using a non-uniform orthogonal grid; applying extrapolation procedures to time signatures; and, using distributed computing techniques. It is possible to achieve significant savings in

computational time and memory requirements through a judicious application of these techniques. This has been demonstrated by analyzing a two-layer stacked patch antenna.

Despite the recent advances in the field of numerical modeling of microstrip antennas, it is not always possible to predict the behavior of an arbitrary microstrip structure accurately and efficiently and the modeling of complex geometries still remains a relatively time consuming process. This is especially true for certain antennas, including the previously analyzed two-layer stacked patch, which often require extensive experimental tuning to obtain the desired impedance and radiation characteristics. However, we believe that this gap will be bridged with further advances in the computing power of future platforms. In any event, when we are investigating a particular class of antennas, we can choose a suitable modeling tool to analyze a few representative cases and employ interpolation techniques to develop CAD formulae [26] which can then be used to predict the trends in the antenna parameters as the characteristics of the antenna are varied.

5. References

[1] D. M. Pozar, "Microstrip Antennas", *Proc. IEEE*, vol. 80, no. 1, pp. 79-91, Jan. 1992.

[2] W. F. Richards, "Microstrip Antennas", Chapter 10, *Antenna Handbook*, Y. T. Lo and S. W. Lee, eds., pp. 10:1-74, New York: Van Nostrand Reinhold, 1988.

[3] W. F. Richards, Y. T. Lo, and D. D. Harrison, "An improved theory for microstrip antennas and applications", *IEEE Trans. Antennas Propagat.*, vol. 29, no. 1, pp. 38-46, Jan. 1981.

[4] J. R. James and P. S. Hall, *Handbook of Microstrip Antennas,* London, U. K.: Peter Peregrinus, 1989.

[5] E. H. Newman and P. Tulyathan, "Analysis of microstrip antennas using moment methods", *IEEE Trans. Antennas Propagat.,* vol. 29, no. 1, pp. 47-53, Jan. 1981.

[6] J. R. Mosig and F. E. Gardiol, "General integral equation formulation for microstrip antennas and scatterers", Proc. Inst. Elec. Eng., vol. 132, Pt. H., pp. 424-432, Dec. 1985.

[7] K. S. Kunz and R. J. Luebbers, *The finite difference time domain method for electromagnetics,* CRC press, 1993.

[8] R. Mittra, W. D. Becker, and P. H. Harms, "A general purpose Maxwell solver for the extraction of equivalent circuits of electronic package components for circuit simulation", *IEEE Trans. Circuits Syst.-I: Fundamental Theory and Applications,* vol. 39, no. 11, pp. 964-973, Nov. 1992.

[9] A. Taflove and K. R. Umashankar, "Review of FD-TD numerical modeling of electromagnetic wave scattering and radar cross section", Proc. IEEE, vol. 77, pp. 682-699, May 1989.

[10] A. Reineix and B. Jecko, "Analysis of microstrip patch antennas using the finite difference time domain method", *IEEE Trans. Antennas Propagat.,* vol. 37, no. 11, pp. 1361-1369, Nov. 1989.

[11] D. M. Sheen, S. M. Ali, M. D. Abouzahra and J. A. Kong, "Application of the three-dimensional finite-difference time domain method to the analysis of planar microstrip circuits", *IEEE Trans. Microwave Theory Tech.,* vol. 38, no. 7, pp. 849-857, July 1990.

[12] E. M. Daniel, C. J. Railton, "Fast finite difference time domain analysis of microstrip patch antennas", *IEEE APS Symposium,* vol. 1, pp. 414-417, June 1991.

[13] C. Wu, K. L. Wu, Z. Bi, J. Litva, "Modeling of coaxial-fed microstrip patch antenna by finite difference time domain method", *Electron. Lett.,* vol. 27, no. 19, pp. 1691-1692, Sept. 1991.

[14] W. J. Buchanan, N. K. Gupta and J. M. Arnold, "Simulation of radiation from a microstrip antenna using three-dimensional finite-difference time-domain (FDTD) method", *Proc. Eighth Int'l Conf. Ant. Propagat.,* IEE Conf. Publ. 370, pp. 639-642, 1993.

[15] B. Toland, J. Lin, B. Houshmand, and T. Itoh, "FDTD analysis of an active antenna", *IEEE Microwave and Guided Wave Letters,* vol. 3, no. 11, Nov. 1993.

[16] K. S. Yee, "Numerical solution of initial boundary value problems involving Maxwell's equations in isotropic media", *IEEE Trans. Antennas Propagat.,* vol. 14, no. 3, pp. 302-307, May 1966.

[17] K. Mahadevan, R. Mittra and P. M. Vaidya, "Use of Whitney's edge and face elements for efficient finite element time domain solution of Maxwell's equations," to appear in the *Journal of Electromagnetic Wave and Applications.*

[18] T. Kashiwa, T. Onishi and I. Fukai, "Analysis of microstrip antennas on a curved surface using conformal grids FD-TD method", *IEEE Trans. Antennas Propagat.,* vol. 42, no. 3, pp. 423-427, March 1994.

[19] Y. Hua and T. K. Sarkar, "Generalized pencil-of-function method for extracting poles of an EM system from its transient response," *IEEE Trans. Antennas Propagat.,* vol. 37, no. 2, pp. 229-234, Feb. 1989.

[20] J. Litva, C. Wu, K. L. Wu, J. Chen, "Some considerations for using the finite difference time domain technique to analyze microwave integrated circuits," *IEEE Microwave and Guided Wave Letters,* vol. 3, no. 12, pp. 438-440, Dec. 1993.

[21] R. J. Luebbers, K. S. Kunz, M. Schneider, F. Hunsberger, "A finite-difference time-domain near zone to far zone transformation", *IEEE Trans. Antennas Propagat.,* vol. 39, no. 4, pp. 429-433, April 1991.

[22] V. Varadarajan and R. Mittra, "Finite-Difference Time-Domain (FDTD) analysis using distributed computing", *IEEE Microwave and Guided Wave Letters,* vol. 4, no. 5, pp. 144-145, May 1994.

[23] K. F. Lee, S. Chebolu, W. Chen, and R. Q. Lee, "On the role of substrate loss tangent in the cavity model theory of microstrip patch antennas", *IEEE Trans. Antennas Propagat.,* vol. 42, no. 1, pp. 110-112, Jan. 1994.

[24] R. Yang, R. Mittra and M. Itoh, "A new omnidirectional CP patch antenna", *IEEE APS Symposium,* vol. 3, pp. 1848-1851, June 1994.

[25] G. Mur, "Absorbing boundary conditions for the finite-difference approximation of the time-domain electromagnetic field equations," *IEEE Trans. Electromagn. Compat.,* vol. 23, pp. 1073-1077, Nov. 1981.

[26] W. Chen, K. F. Lee, J. S. Dahele and R. Q. Lee, "CAD formula for the resonant frequency of rectangular patch antenna with dielectric cover", *Proc. Eighth Int'l Conf. Ant. Propagat.*, IEE Conf. Publ. 370, pp. 550-553, 1993.
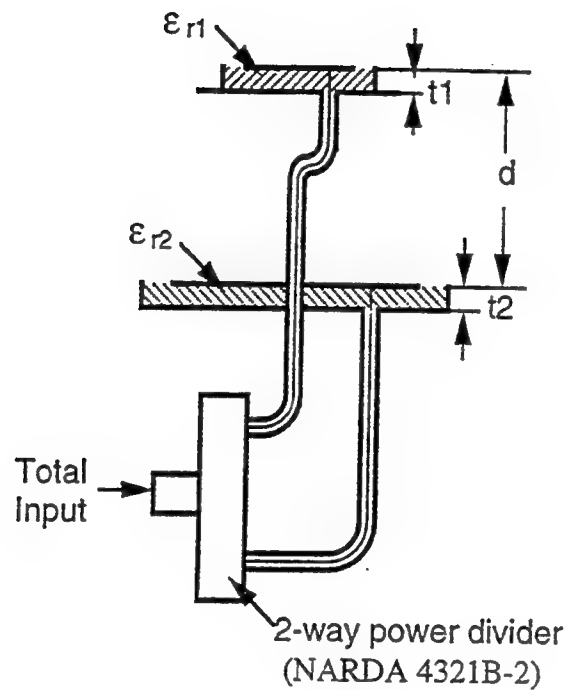
$d = 71$mm $\varepsilon_{r1} = 10.5$
$t_1 = t_2 = 3.12$mm $\varepsilon_{r2} = 3.6$

$a_1 = 29.3$mm, $b_1 = 28.8$mm, $g_1 = 43.6$mm
$a_2 = 50.9$mm, $b_2 = 48.9$mm, $g_2 = 60.8$mm

a, b---Dimensions of patches
g---Dimensions of ground planes

Fig. 1 Two-layer stacked patch antenna.

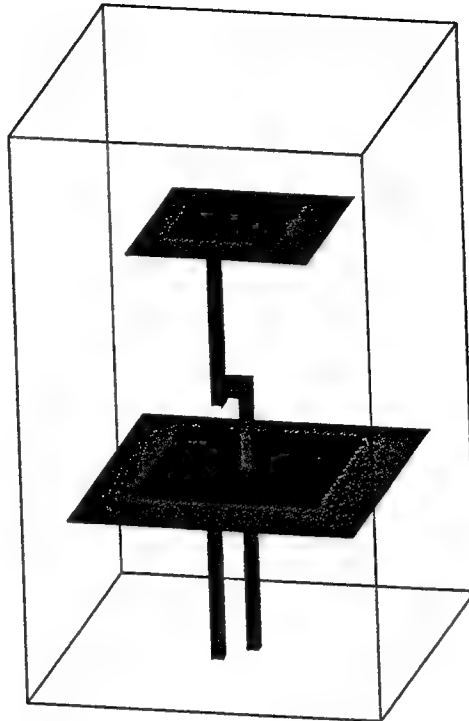Fig. 2(a): Discretized geometry
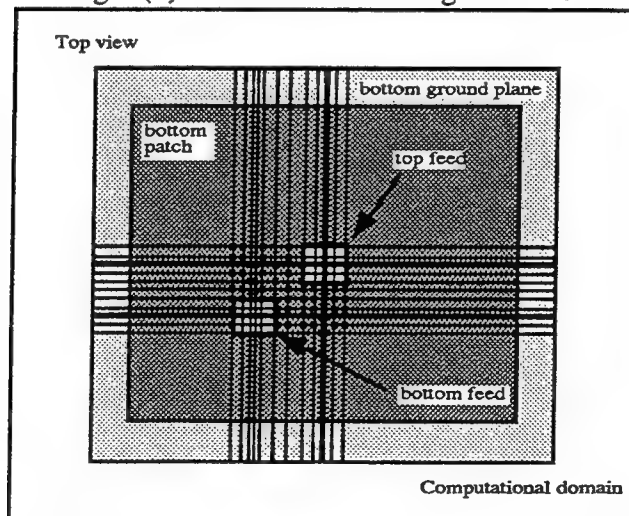

Fig. 2(b): Non-uniform orthogonal mesh

Fig. 2 FDTD modeling of the two-layer stacked patch antenna.
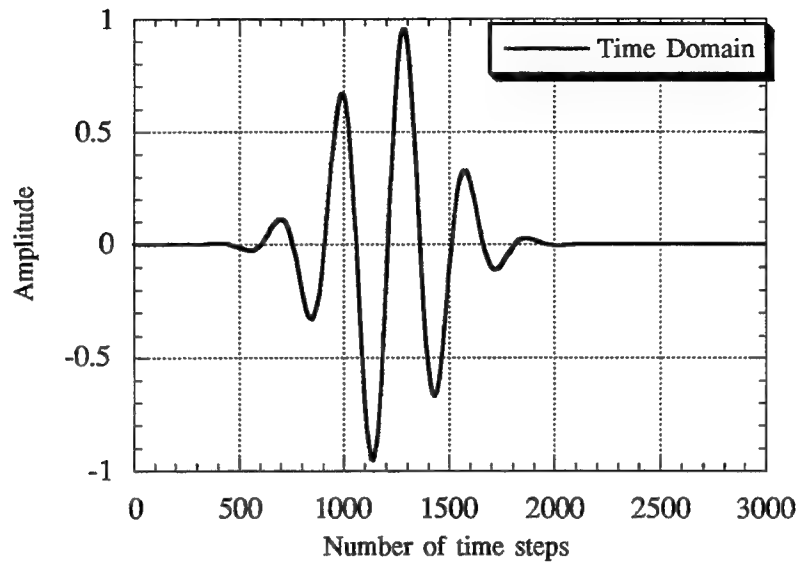
Fig. 3 (a)

Fig. 3 (b)
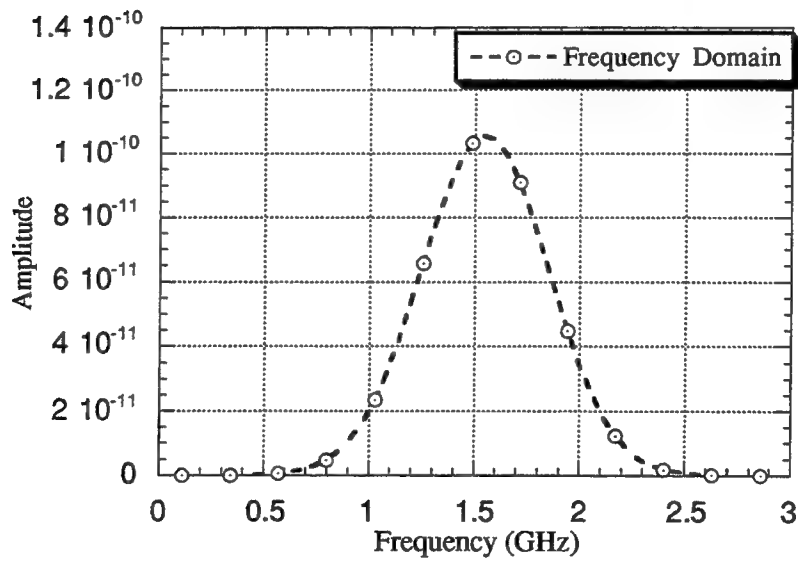
Fig. 3 Input Excitation: A 1.55 GHz sinusoid modulated with a Gaussian

pulse having a 3dB cutoff frequency of 0.25 GHz.

Fig. 4(a)

Fig. 4(b)
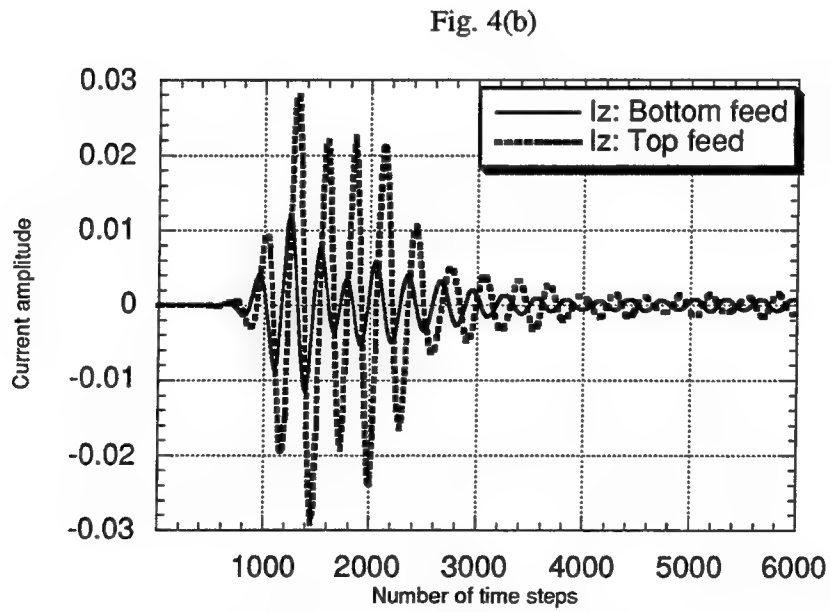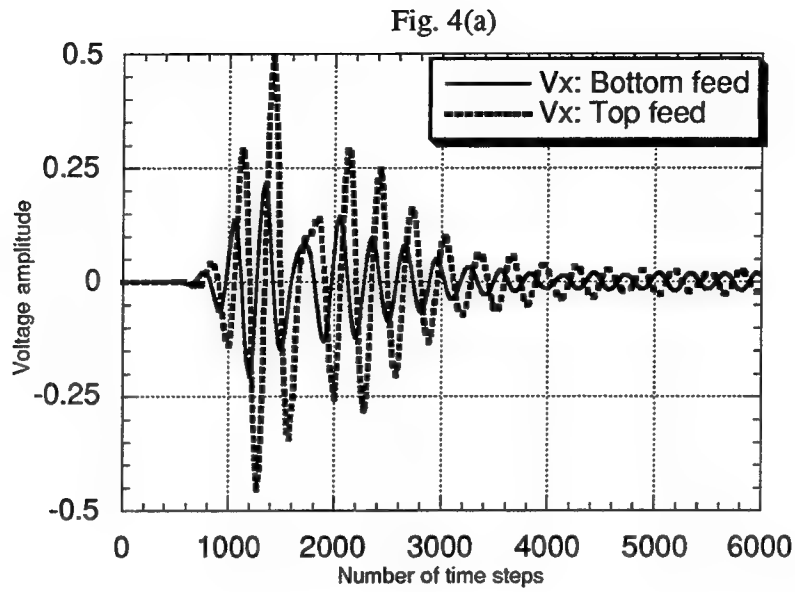
Fig. 4 Time domain voltage and current waveforms

304
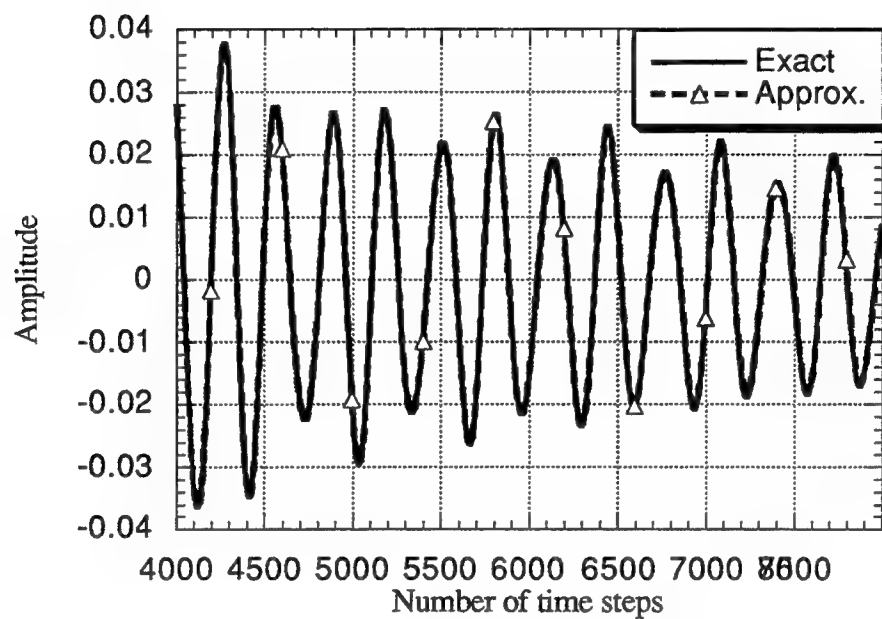
Fig. 5 Results of extrapolating of time-domain voltage waveform for the top patch.

     _____   Data obtained from the FDTD simulation.
    --Δ--   Extrapolated waveform using GPOF algorithm with interpolation between 4000-6000 samples.

Fig. 6(a): Real and imaginary parts of the input impedance, $Z_{in} = R_{in} + j\,X_{in}$.

Fig. 6(b): Normalized reflection coefficient.

Fig. 6 Impedance results for the bottom patch.

Electric field, Ez at 7000th time step

Current density, Jx at 7000th time step

Currenty Density, Jy at 7000th time step

Fig. 7   Field distributions on the bottom patch

Pulsed Antenna Measurements
With the HP 8530A Microwave Receiver

John Swanstrom - Applications Engineer
and
Robert Shoulders - Design and Development Engineer

Hewlett-Packard Company
Santa Rosa Systems Division
1400 Fountaingrove Parkway
Santa Rosa, CA 95403

Abstract

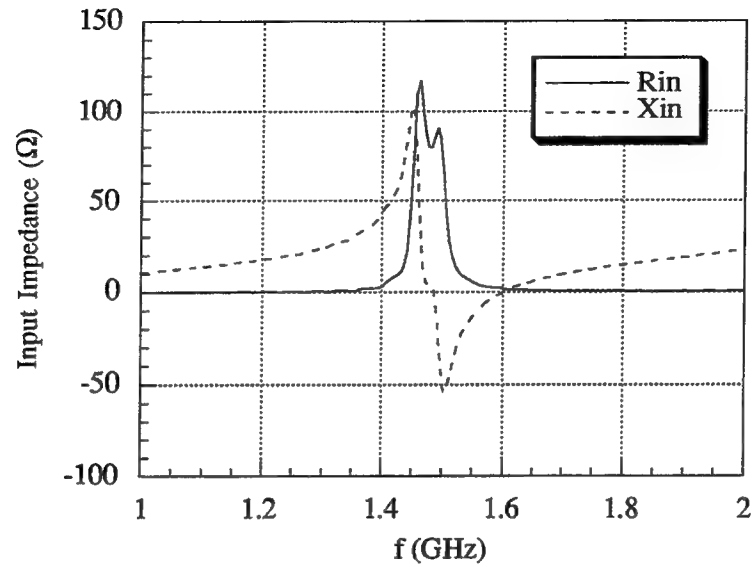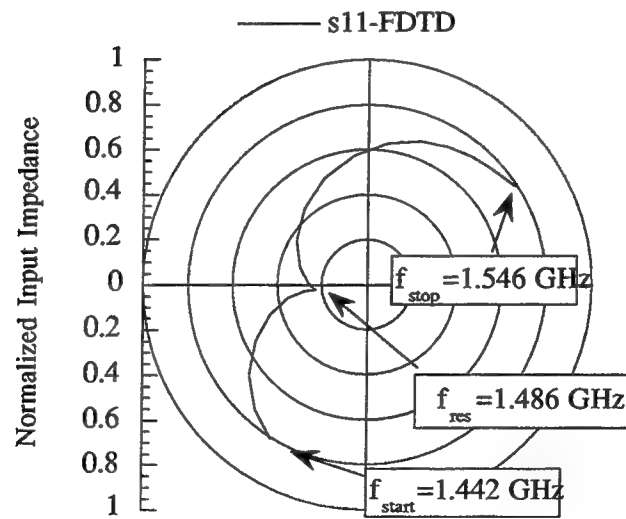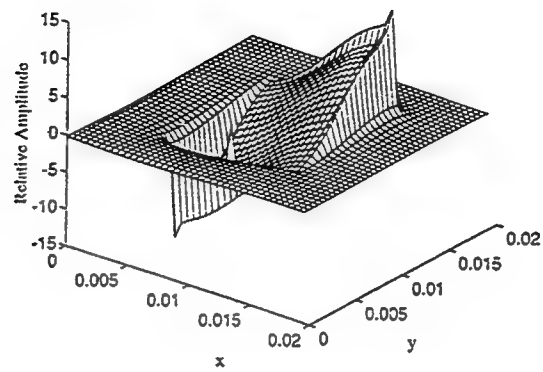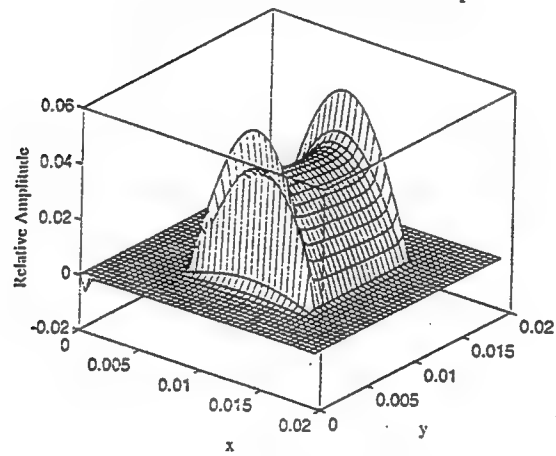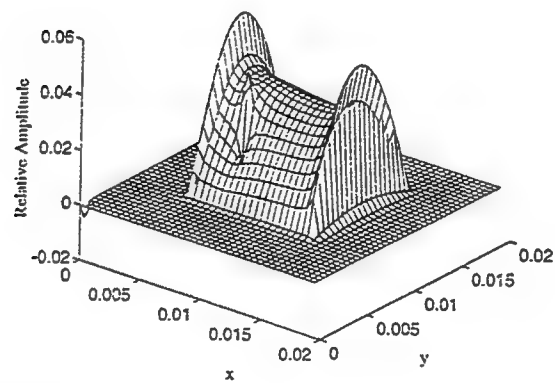This paper discusses the instrumentation techniques that can be used for the measurement and characterization of antennas that are to be tested in a pulsed-RF mode of operation. A pulse-parameter chart is presented that illustrates all possible ranges of pulse width and pulse repetition frequencies for antennas operating in a pulsed mode. An antenna operating in a pulse mode will have pulse parameters that lie somewhere on the pulse-parameter chart. This paper defines five different measurement regions of the pulse-parameter chart, and presents the measurement techniques for measuring pulsed antennas that operate in each of these regions.

Keywords: Pulsed antenna measurements, instrumentation, HP 8530A microwave receiver

1.0 Introduction

Many of today's high-performance antennas utilize solid-state transmit/receive (T/R) modules which give antenna designers the building blocks to design high-performance active-array antennas. Often these active-array antennas are designed to operate in a pulsed mode. When an antenna is utilized in a pulsed application, it is usually desirable to test it under the same pulsed conditions. This paper presents several different measurement techniques that can be used to measure antennas operating in a pulsed mode, and will show that the appropriate technique will depend upon the pulse parameters of the antenna and the characteristics of the measurement receiver. It will be shown that measurement sensitivity is one of the key factors that affect the performance of a pulsed measurement system.

2.0 General Concepts for Pulsed Measurements

Pulsed antenna measurements can be described in terms of the pulse width, which is the duration of a single pulse, and the pulse repetition frequency (PRF). The pulse period is the reciprocal of the PRF. The duty cycle is defined as the ratio of the pulse width to the pulse period.

Receiver characteristics which are of particular importance to pulsed measurements are:
- rise time -- the time required for the receiver response to rise to $1/\varepsilon$ times the steady state response of a stepped RF signal,
- stop bandwidth -- bandwidth around the measurement frequency beyond which the system response is essentially null,
- system settling time (T_S) -- the time required for the system to settle to within 0.1 dB to a stepped RF signal.

Receiver performance in pulsed measurements is a function of both the pulse parameters and these receiver parameters. In general, for a given receiver there are five regions of operation. These regions, CW, high PRF, full pulse characterization, peak

response, and excluded region, can be graphed on a plot of pulse period vs. pulse width, as shown in Figure 1. The boundaries between these regions and the system performance within them will be discussed individually below, using the two IF bandwidths of the HP 8530A Option 008 microwave receiver as examples.
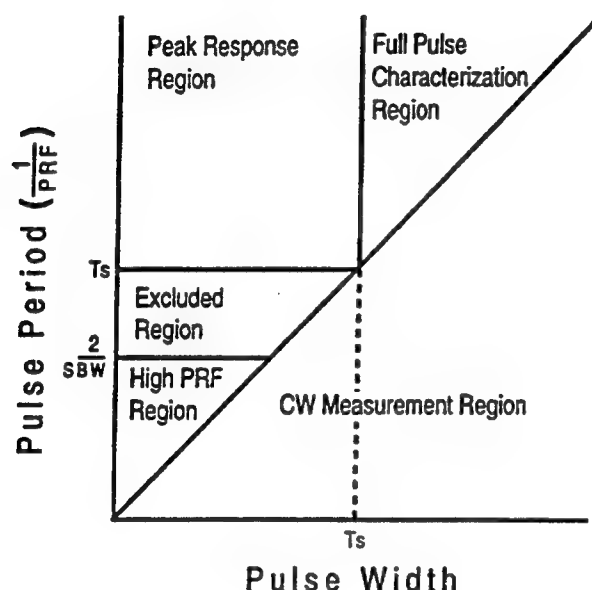

Figure 1: General pulse-parameter chart for receivers operating in a pulsed mode.

3.0 HP 8530A Option 008 Pulsed Microwave Receiver

The HP 8530A Option 008 offers two independent IF bandwidths, each with unique pulse characteristics.

The narrow IF bandwidth (standard on the HP 8530A) has a rise time of 60 us and a system settling time (T_s) of 180 us. A unique attribute of this IF is its narrow stop bandwidth of 70 KHz, which makes it particularly well suited for operation in the high PRF region, as explained below.

The wide IF bandwidth, which is added to the HP 8530A with Option 008, has a 300 ns rise time, and a system settling time (T_s) of 1 us. This makes it especially useful for full-pulse characterization.

Accurate measurements can be made for virtually any combination of pulse width and pulse period encountered in pulsed antenna measurements by utilizing these two different IF bandwidths.

4.0 Pulse Measurement Regions

4.1 CW Measurement Region

In the region where the pulse width exceeds the pulse period, the pulse is always "on" and the antenna is said to be operating in a CW measurement mode. This is a common operating condition for many antennas and the measurement techniques are well known.

4.2 High PRF Region

In the region of the chart where the PRF exceeds 1/2 the receiver stop bandwidth, the pulse harmonics are eliminated by the receiver filter leaving only the center frequency. As a result, the receiver responds to the high PRF-pulsed signal as though it were a CW signal [1]. The effect of the high PRF pulsing on the measurement is a reduction in the signal level measured by the receiver due to the loss of the energy in the pulse harmonics. This loss of energy is known as duty-cycle pulse desensitization which will reduce the sensitivity of the measurement system by approximately 20 log [duty cycle]. Highly sensitive measurement systems can usually tolerate some desensitization loss. With receivers that have less sensitivity, averaging can always be used to recover sensitivity if necessary, but at a penalty of decreased measurement speed. Thus selecting a receiver with the best sensitivity available will maximize the accuracy, dynamic range, and measurement speed in the high PRF region.

Note that the loss of sensitivity depends on the duty cycle, and not the pulse width, of the measurement. Thus narrow-bandwidth receivers can make good measurements of extremely narrow pulses as long as the duty cycle is sufficiently high and the PRF is greater than 1/2 the stop bandwidth. For the HP 8530A, the high PRF region begins at a PRF of 35 KHz. Figure 2 shows the sensitivity of the HP 8530A in the high PRF region.
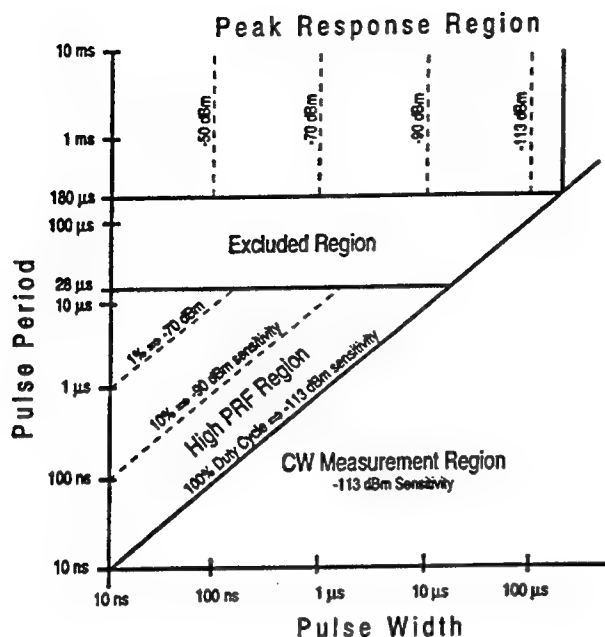
Figure 2: HP 8530A receiver sensitivity (using narrow band IF) showing duty cycle desensitization in the high PRF region, and pulse width desensitization in the peak response region.

4.3 Full Pulse Characterization Region

If the pulse width exceeds the receiver settling time, then the measurement falls into the full pulse characterization region (see Figure 1). In this region, the receiver performs at its full CW sensitivity and dynamic range with no pulse desensitization. A unique attribute of measurements in this region is that the antenna being measured may be characterized within the pulse to a resolution equal to the rise time of the receiver. This makes possible two types of analysis, pulse profile and point-on-pulse, both of which are available with the HP 8530A Option 008. With the shorter rise time and faster system settling time (T_s) of this option, the HP 8530A is able to perform pulse characterization utilizing either of these modes on pulses as short as 1 us with a resolution of 300 ns.

Pulse Profile Mode of Operation

The pulse-profile mode of operation enables the receiver to profile the pulse out of the antenna-under-test (AUT) and display the pulse on the receiver's CRT display. This pulse-profile mode provides a powerful analysis capability enabling the antenna designer to view the pulse out of the AUT and look for characteristics such as pulse droop. Pulse-profile measurements have been made over the full frequency range (0.1-110 GHz) with the HP 85301B antenna measurement system.
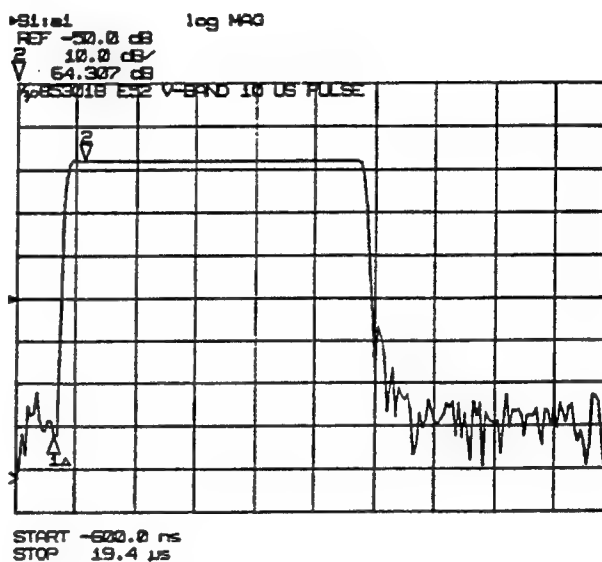
Figure 3: Pulse profile of a 10 microsecond pulse displayed on the microwave receiver's display, 50 GHz operating frequency, 201 data points, no averaging applied.

Point-on-Pulse Mode

With the new point-on-pulse mode of operation, an antenna designer now has the ability to measure an antenna pattern at any point on the pulse, using the pulse trigger delay. This allows comparing antenna patterns from different points of time in the pulse.

310

4.4 Peak Response Region

If the pulse width is less than the system settling time (T_S) and the pulse period is greater than T_S, then the measurement falls in the peak response region of operation (see Figure 1). In this region, the system is partially charged to a stable value by each individual pulse. This partial charging results in pulse desensitization (a loss of dynamic range and sensitivity) of approximately 20 log [pulse width/rise time]. Notice that, unlike the high PRF region in which the pulse desensitization depends on the duty cycle, this desensitization depends on the pulse width alone, which sets a practical limit on the pulse width that can be measured. Figures 2 and 4 show the sensitivity of both bandwidths of the HP 8530A microwave receiver in the peak response region of the pulse-parameter chart.
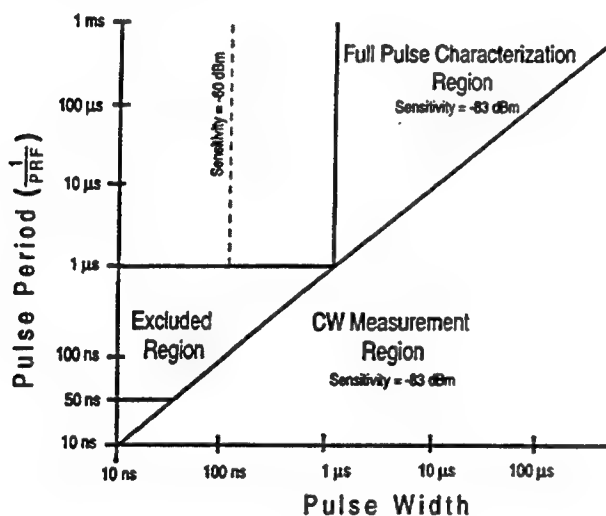
possible, the required timing complexity makes measurements in the excluded region undesirable. By utilizing two complementary IF bandwidths, the peak response region of the wide band IF overlaps the high PRF region of the narrow band IF effectively eliminating the excluded region as shown in Figure 5.

Note that for PRFs between 35 KHz and 1 MHz, the HP 8530A can be operated either in the high PRF region using the narrow-band mode or in the peak response region using the wide-band mode. The preferred method will usually depend on which approach yields the highest sensitivity. By utilizing two IF bandwidths, HP has been able to overcome the limitations of the excluded region, and allow measurement of pulsed antennas in any region of the pulse-parameter chart.

Figure 4: HP 8530A receiver sensitivity with the wide bandwidth IF/detectors.
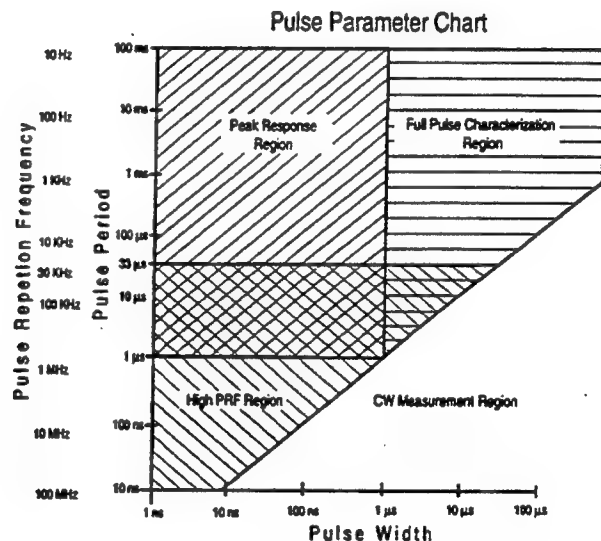
Figure 5: Pulse parameter chart showing operating regions for the HP 8530A Option 008 microwave receiver. Note that with two IF bandwidths there is no excluded region.

4.5 Excluded Region

For any given IF bandwidth, there will always be an excluded region in the pulse-parameter chart (see Figure 1). The excluded region is bounded by pulse periods less than the receiver settling time (T_S), and greater than 1/[2*receiver stop bandwidth]. In this region, the receiver's response is a slowly stabilizing train of pulses with increasing peak levels. Although measurement of this type of response is

5.0 AUT in High Power Transmit Mode

Many active element phased array antennas are used in pulsed radars, and in operation they will transmit in a high-power mode and receive in a low-power mode. To verify proper operation, it will be necessary to test the antenna's patterns in both

311

modes of operation. Testing in both modes requires a dual-directional antenna range that will operate in a pulsed mode. Figure 6 illustrates a typical dual-directional antenna range that can measure pulsed antennas with any pulse width and pulse period combination on the pulse-parameter chart . The transmit configuration cannot use a radiated reference since it would interfere with the signal from the (transmitting) AUT. Therefore, a low-loss cable is used to provide the reference signal. When this is not practical, scalar-only measurements are still possible because the system's synthesized local oscillator eliminates the need for a phase-lock reference signal. The measurement technique selected will depend upon where the antenna operates on the pulse-parameter chart. Variations to the measurement configuration can be designed to meet unique measurement requirements.

System Sensitivity

As with all antenna measurements, sensitivity of the measurement system is very important for measurement performance. Sensitivity determines the minimum signal that can be detected by the measurement system, and is an important issue on antenna ranges which always have large amounts of dispersive losses. Pulse measurements also place an additional demand on sensitivity due to duty cycle and pulse desensitization loss. Sensitivity also affects the signal-to-noise ratio which is usually the largest factor in determining measurement accuracy. Finally, sensitivity can affect the speed of a measurement. If insufficient sensitivity is available to achieve the measurement accuracy required, averaging can be used to improve the sensitivity of the measurement system, but at a penalty of measurement speed.

Summary

While pulsed operating conditions add some unique challenges to antenna measurements, it was shown that pulsed-antenna measurements can be easily performed with only slight modifications to a standard CW antenna measurement system. Three different techniques were presented for measuring pulsed antennas. Identifying the region of the pulse-parameter chart in which the pulsed antenna operates determines which of these measurement techniques to use.

Conclusions

The use of two different measurement bandwidths in the microwave receiver allows measurement of pulsed antennas for all pulse widths and pulse periods. Narrow bandwidths (with their inherent sensitivity) are optimum for making CW or high-PRF pulsed-antenna measurements. For full-pulse characterization, wide bandwidth receivers provide better resolution for point-on-pulse and pulse profile analysis. In addition, receiver operation in the peak response region allows measurements to be made when narrow pulses and low PRFs are required.

References

[1] Product Note 8510-9; Pulsed-RF Network Measurements Using the HP 8510B Network Analyzer, by Hewlett-Packard.
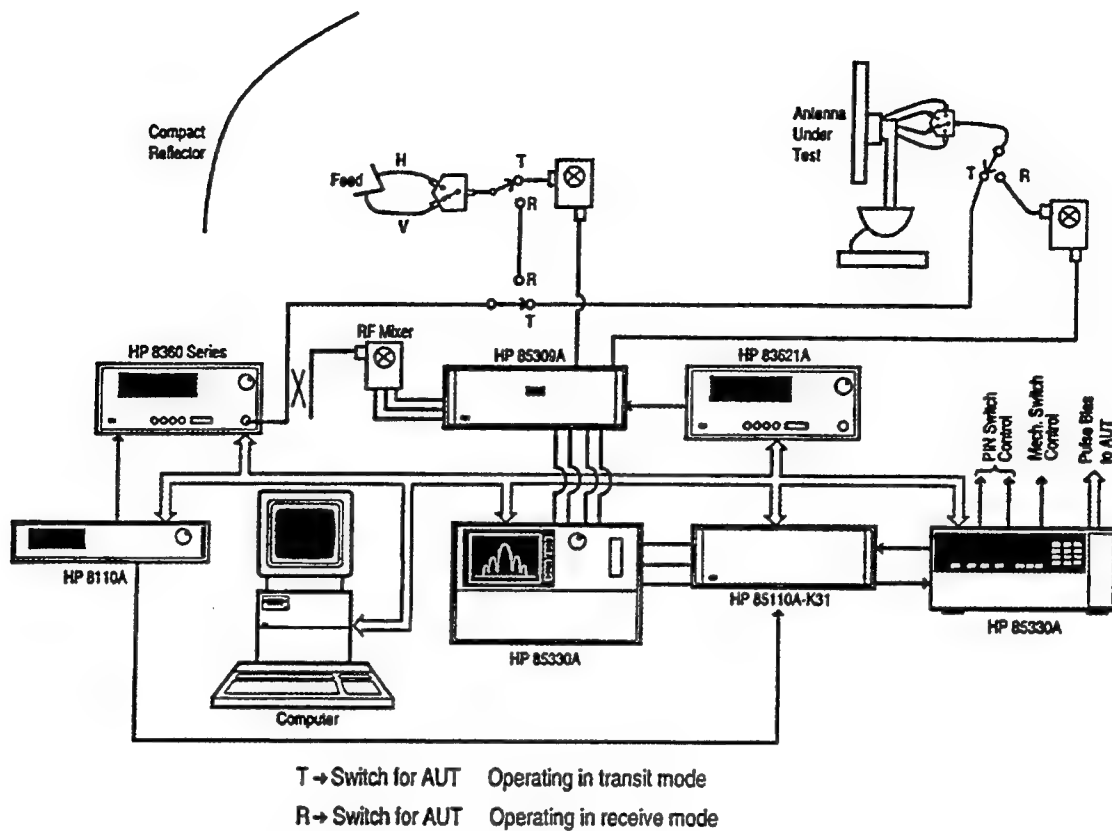
Figure 6: Dual-directional antenna range which can measure pulsed antennas in either a transmit or receive mode in all regions of the pulse-parameter chart.

RESONANCES IN LARGE CIRCULAR ARRAYS:
AN EXPERIMENTAL STUDY

George Fikioris
Rome Laboratory
Electromagnetics and Reliability Directorate
Hanscom AFB, MA, 01731

Abstract: Recent theoretical studies [1] predict that large circular arrays of cylindrical dipoles should have very narrow resonances as the frequency is varied. Only one element is driven, and the rest are parasitic. At each resonance, the driving-point conductance is large, and the driving-point susceptance is zero. An experimental study of this phenomenon is described in the present paper. The experimental circular array consists of 90 vertical brass monopoles over an aluminum ground plane. The monopoles are electrically short and thick. The single driven element is the extension of the inner conductor of a coaxial transmission line. The driving-point admittance of the array is determined from swept-frequency measurements of voltages (magnitudes and phases) at fixed points on the coaxial line. Difficulties inherent in the determination of the rapidly varying driving-point admittance are discussed. Narrow resonances are found when the frequency is close to 2.5 GHz. The resonant frequencies obtained experimentally are compared to those obtained from the theory, and the agreement is excellent.

1. Introduction and background

The phenomenon of resonances in large circular antenna arrays has been studied extensively in the past few years, both theoretically and experimentally. The idea that a large closed-loop array may possess extremely narrow resonances came from studies in quantum mechanics [2,3] as well as from theoretical and experimental studies conducted in the past with arrays of cylindrical dipoles. These initial considerations were reported by R. W. P. King in [4]; this paper also mentioned the important possible connection between resonance and superdirectivity in a large closed-loop array.

The "two-term theory" [5] is an approximate solution of the integral equations for the current distributions on the dipoles. As a first step in the investigation of resonances in a large circular array, numerical calculations were performed using the two-term theory in its original form. These numerical calculations led to the discovery of narrow resonances in $N =$60, 72,

and 90- element arrays when the half-length h and radius a of the dipoles are $h/\lambda = 0.18$ and $a/\lambda = 0.028$. Only one of the elements in the array is driven, and the rest are parasitic. The resonances are illustrated graphically in [6]. In the Figures of [6], the spacing d/λ between adjacent elements is the varying parameter; its value is very close to 0.3 in the region in which narrow resonances occur. Each resonance corresponds to a maximum in the driving-point conductance of the array as d/λ is varied. Each resonance is characterized by a large integer parameter m, with $m \leq N/2$. Resonances were discovered for this rather unusual combination of h/λ and a/λ after unsuccessful attempts were made to discover resonances in arrays with combinations commonly used in other applications.

A $N = 90$-element experiment [7] with monopoles over a ground plane was performed with the parameters of [6] and narrow resonances were observed experimentally. In the experimental study, the parameters a, h, and d are fixed, and the frequency is varied close to 2 GHz.

In a subsequent paper, Freeman and Wu re-examined the kernels used in the integral equations and derived a new set of kernels [8,9]. These kernels were incorporated into the two-term theory, and a detailed analytical/numerical study of the modified two-term theory formulas led to the development of guidelines for the choice of the parameters a, h, and d [10, Ch. 4], [1]. In these references, the current distributions on the parasitic elements and the associated field patterns of the resonant array are also discussed. It then became possible to choose a, h, and d so that the resonances are much narrower than those presented in the Figures of [6]. Although the modified theory in general predicts much narrower resonances than those of the original theory, this is not true for the specific parameters of the numerical calculations in [6] and used subsequently for the design of the experiment in [7].

In this paper, an experimental study of a circular array in which the parameters a, h and d are chosen based on the modified theory contained

in [1] is presented. The present experimental study is described in more detail in [10, Ch. 8]. The resonances are theoretically predicted to be very narrow. The narrowness of the peaks as well as the high SWR in the feeding-measuring transmission line are the two main complicating factors in the experiment.

2. Layout and choice of parameters

The general layout is shown in Figures 1a and 1b and is very similar to the one of [7]. The parasitic elements are vertical, thin-walled brass tubes (monopoles) fastened onto a horizontal aluminum ground plane. The driven element is the extension of the inner conductor of a coaxial transmission line, so that the driving-point admittance of the array is the terminating admittance of the transmission line. The purpose is to calculate the driving-point admittance $Y_{1,1}(f) = G_{1,1}(f) + jB_{1,1}(f)$ of the array from swept-frequency measurements of voltages (magnitudes and phases) at fixed points on the transmission line. The Hewlett-Packard HP8753C Network Analyzer is to be used as a swept-frequency generator and as a swept-frequency vector voltmeter. Because of its high stability and fine resolution, it is well-suited for measurements involving narrow resonances.

Narrow resonances in large circular arrays of electrically short elements exist only if the array parameters (number of elements N, monopole length h, monopole diameter $2a$, and array diameter $2R$) are properly chosen. In this experiment, the parameters were chosen to be

$$N = 90, \quad h = 0.858 \text{ in.}, \quad 2a = 1/4 \text{ in.}, \quad 2R = 40 \text{ in.} \tag{1}$$

With these parameters, the theoretically predicted resonant frequencies are

$$f_{33} = 2.55 \text{ GHz}, \quad f_{45} = 2.68 \text{ GHz}, \tag{2}$$

for the $m = 33$ and the $m = 45$ (last) resonance, respectively. Both the physical size of the array and the resulting operating frequencies were suitable

316

for measurements with the available equipment and facilities. The resonant frequencies in (2) were obtained using the theory of [1], including the refinements in Section 6 of [1]. The theoretically predicted electrical parameters at the $m = 33$ and $m = 45$ resonances are

$$h/\lambda_{33} = 0.185, \ a/\lambda_{33} = 0.027, \ d/\lambda_{33} = 0.301; \tag{4}$$

$$h/\lambda_{45} = 0.195, \ a/\lambda_{45} = 0.028, \ d/\lambda_{45} = 0.317, \tag{5}$$

respectively. Such an array is predicted to yield extremely narrow resonances if the elements are perfectly conducting. Extremely narrow resonances can be realized in practice only with superconducting elements; in the case of brass elements over an aluminum ground plane, the narrowness of the peaks is severely limited.

Note that the relative frequency range of interest,

$$(\lambda_{33} - \lambda_{45})/\overline{\lambda} = (f_{45} - f_{33})/\overline{f} \simeq 0.05, \tag{6}$$

is very small. This smallness simplifies the design of the transmission line; its dimensions and parameters were chosen assuming a single operating wavelength equal to $\overline{\lambda} = (\lambda_{33} + \lambda_{45})/2 = 4.52$ in.

3. Constructional details

The main requirement of the construction is precise mechanical implementation. For example, the elements must be perfectly vertical, and positioned precisely on the circle. The need for precise implementation is obvious from the nature of the phenomenon (very narrow resonances), but a beforehand theoretical estimate of the effects of imperfect mechanical implementation was not available. For example, to estimate the necessary accuracy in positioning the elements, an analysis of a noncircular array would have to be

317

performed. Although this is possible in principal with the theory in [10, Ch. 2], it is difficult numerically. This is why the mechanical construction was made, within a reasonable cost, as precise as possible. Another general consideration in the design was to have good electrical contact between the brass monopoles and the aluminum ground plane. This was shown to be very important in the experimental study of [7].

The design for the 89 parasitic monopoles is shown in Figure 2. The walls were made quite thin (the wall thickness is $a - a_1 = 1/8 - 7/32 = 1/64$ in.), so that the tubular elements of the theoretical model could be realized. The tolerance in the height of the monopoles is ± 0.001 in. Each one was fastened with a screw onto a 1/4 in. thick jigplate aluminum sheet measuring 60 in. by 60 in.. This design provides good electrical contact between monopoles and ground plane. The smallest distance between the array and the ground-plane edge is about 2.2 predicted average wavelengths $\bar{\lambda}$. The tolerances in the location of the elements (tolerances of the 89 tapped holes in the ground plane) was ± 0.001 in. in the radial positioning and $\pm 0°0'1''$ in the angular positioning. The other (90th) hole measured 3/4 in. in diameter, the size of the outer conductor of the transmission line.

The coaxial transmission line is shown in Figure 3. It is made of brass, has an outside diameter of $2b = 0.687$ in., and an inside diameter of $2a = 1/4$ in., the same as the parasitic element diameter. The characteristic impedance is $Z_c = 60 \ln(2b/2a) = 60$ Ohms and the characteristic admittance is $Y_c = Z_c^{-1} = 17$ mA/V. This coaxial line is a single mode, low-loss transmission line. The skin depth $d_s = (\pi f \mu_0 \sigma_B)^{-1/2}$ of brass is of the order of 10^{-3} mm for the operating frequencies of interest. If the propagation constant of the TEM mode is calculated, it is seen that it is larger than the attenuation constant by a factor of about 4000.

The inner conductor in Figure 3 is supported by two brass shorting plugs (rings) underneath the feeding line (probe). The inner conductor is also supported by a thin ring of styrofoam located about 1 in. from the

top of inner conductor (top of the driven monopole). When measurements were taken, it was found that the values of the driving-point conductance at resonance were very sensitive to adjustments in the positioning of the transmission line-driven element; special care was required in order to position the inner conductor correctly.

Four 1/4 in. holes were drilled on the outer conductor of the transmission line so that the same voltage probe could move from hole to hole. The reason for using one voltage probe is that it is difficult to construct two identical probes. The distances z_A, z_B, z_C and z_D from the top of the transmission line to each hole (Figure 3) were chosen so that 1) the distances between holes would not be an integral multiple of $\overline{\lambda}/2$ at the frequency interval of interest. If this were the case, the voltage probe would measure the same or opposite voltages at the two holes; 2) Any of the six (measurements are taken at two points) possible choices of probe locations would result in a different electrical distance. The distances between holes in predicted average wavelengths is roughly $z_B - z_A = 0.65\overline{\lambda}$, $z_C - z_A = 1.36\overline{\lambda}$, $z_D - z_A = 1.93\overline{\lambda}$, $z_C - z_B = 0.71\overline{\lambda}$, $z_D - z_B = 1.28\overline{\lambda}$, and $z_D - z_C = 0.57\overline{\lambda}$, so that the closest integer multiple of $\overline{\lambda}/2$ occurs for the probe locations AD and CD.

Standard cables and connectors were used to connect the generator to the brass transmission line and the voltage probe to the vector voltmeter. No provision was made for matching the feeding coaxial cables to the brass transmission line since the rapidly varying load impedance of the brass line (driving-point impedance of the array) would make this impossible. This does not matter for the experiment as long as enough power is supplied to the load (array) because the method of measurement does not depend on the conditions of the generator. Such matching, however, would require special care in an application in which a band of frequency near a narrow resonance is of interest.

4. Measurement method

The measurement method with which the results that follow were obtained is the "fixed-point method" developed in [7] for a lossless transmission line. Here, the ohmic losses in the line are also taken into account. It is a fixed-point, swept-frequency method in the following sense: Measurements of voltages are taken at a fixed point on the transmission line (a particular hole in Figure 3) for all frequencies. Then the probe is moved to another point on the transmission line to measure a new series of voltages for the same set of frequencies. This method allows rapid measurements for thousands of frequencies, which are necessary in order to obtain a frequency response curve that shows the resonant admittance values and resonance widths accurately.

The formula relating the voltages V_1 and V_2 at two fixed points along a lossy uniform transmission line to the terminating admittance $Y_L(f)$ of the line is easily obtained using transmission line theory. If z_1 and z_2 are the distances from the load to the two points and a $e^{j\omega t} = e^{j2\pi ft}$ dependence is assumed, $Y_L(f)$ is given by

$$Y_L(f) = Y_c \frac{\cosh \gamma z_2 - (V_2/V_1)\cosh \gamma z_1}{\sinh \gamma z_2 - (V_2/V_1)\sinh \gamma z_1}. \tag{7}$$

In our case, the terminating admittance $Y_L(f)$ is the driving-point admittance $Y_{1,1}(f)$ of the circular array, and z_1, z_2 are z_A, z_B, z_C, or z_D. γ is the complex propagation constant of the line, defined by

$$\gamma = \sqrt{(R + j2\pi fL)j2\pi fC}. \tag{8}$$

R is the resistance per unit length of the transmission line due to ohmic losses on the brass walls of the coaxial line and is given by

$$R = \frac{1}{\pi d_S \sigma_B}\left[\frac{1}{2a} + \frac{1}{2b}\right], \tag{9}$$

where σ_B and $d_S = (\pi f \mu_0 \sigma_B)^{-1/2}$ are the conductivity and skin depth of the brass walls. L and C are, respectively, the inductance and capacitance

per unit length of the line,

$$L = \frac{\mu_0}{2\pi} \ln \frac{b}{a}, \qquad C = \frac{2\pi\epsilon_0}{\ln \frac{b}{a}}. \tag{10}$$

The characteristic admittance Y_c when the ohmic losses in the line are taken into account has a small imaginary part and is given by

$$Z_c = \frac{1}{Y_c} = \sqrt{\frac{R + j2\pi fL}{j2\pi fC}}. \tag{11}$$

In [10, Ch. 8], the sensitivity of the measurement method due to imprecise knowledge of the various quantities appearing in (7)–(11) is discussed. The method is particularly sensitive when the load admittance is large, as occurs when the array is at or near a resonance. However, the important advantage of the method is the fixed point property.

5. Experimental results

Figure 4 shows the experimentally determined frequency response curve for the driving-point conductance $G_{1,1}(f)$ for the probe locations AC. Narrow resonances are seen to occur. The resonant frequencies are those where the driving-point conductance attains its maximum. [10, Ch. 8] contains frequency response curves obtained by other probe locations (AB, BC, and BD). These are similar to the one of Figure 4. In particular, the experimentally determined resonant frequencies are very close for the four combinations; typically, they agree to 4–5 significant digits.

Some differences exist, however, in the measured conductance values at the resonances. These differences are shown in Table 1. The values of resonant driving-point conductance were taken from the data of Figure 4 and the corresponding Figures of [10] for the rest of the probe locations as the maxima of the various peaks. The last row in Table 1 corresponds to the rightmost peak in Figure 4. This peak is identified with both the $m = 44$

m	$G_{1,1}^{AB}$ (mA/V)	$G_{1,1}^{AC}$ (mA/V)	$G_{1,1}^{BC}$ (mA/V)	$G_{1,1}^{BD}$ (mA/V)	$G_{1,1}^{E}$ (mA/V)
29	31.7	30.5	31.4	31.1	31.2
30	51.2	49.5	51.6	51.8	51.0
31	91.3	86.8	94.9	94.7	91.9
32	175.	168.	186.	181.	177.
33	312.	301.	333.	323.	317.
34	480.	467.	491.	478.	479.
35	678.	658.	634.	606.	644.
36	687.	663.	678.	654.	671.
37	629.	607.	625.	617.	619.
38	641.	622.	652.	661.	644.
39	701.	670.	646.	661.	670.
40	702.	692.	720.	738.	713.
41	688.	653.	623.	668.	658.
42	720.	679.	649.	700.	687.
43	749.	693.	582.	692.	679.
44–45	776.	660.	451.	561.	612.

Table 1. Experimentally Determined Values of Resonant Driving-Point Conductance $G_{1,1}$ for Probe Locations AB, AC, BC and BD, and Average Value $G_{1,1}^{E}$.

and $m = 45$ resonances. The coincidence of the two resonant peaks is due to the ohmic losses on the brass elements and the aluminum ground plane and is discussed in detail in [10]. If the monopoles and the ground plane were lossless, the $m = 45$ peak would be separate from and much narrower than the $m = 44$ peak. The last column of Table 1 is the average $G_{1,1}^{E}$ of the four previous columns. The differences in the values of Table 1 are

m	f_m^T (GHz)	f_m^E (GHz)	Percentage Error
29	2.4260	2.4311	0.214
30	2.4623	2.4681	0.236
31	2.4950	2.5009	0.237
32	2.5241	2.5298	0.227
33	2.5497	2.5554	0.224
34	2.5722	2.5777	0.215
35	2.5919	2.5970	0.199
36	2.6090	2.6137	0.182
37	2.6238	2.6288	0.189
38	2.6365	2.6413	0.183
39	2.6473	2.6519	0.176
40	2.6562	2.6602	0.153
41	2.6634	2.6678	0.165
42	2.6689	2.6728	0.146
43	2.6728	2.6765	0.139
44–45	2.6752	2.6791	0.147

Table 2: Resonant Frequencies f_m: Theoretical (f_m^T), Experimental (f_m^E), and Percentage Error.

attributed to: 1) imperfections in the construction of the transmission line; 2) sensitivity to mechanical adjustments of the driven element-extension of the inner conductor of the transmission line; and 3) the sensitivity of the measurement method mentioned in the previous Section. Note that the largest differences in Table 1 occur at the last (m =44–45) resonance. This was the peak that was found to be more sensitive to the slight mechanical adjustments mentioned above. On the other hand, the resonant frequencies

were not found to be sensitive to such adjustments.

Figure 5 shows the experimentally determined values of driving-point susceptance $B_{1,1}(f)$ for the probe locations AC. As predicted by the theory in [1] for lossless elements, there are zeros in this frequency response curve located very close to the frequencies f_m of maximum driving-point conductance. $B_{1,1}(f)$ was also measured with the remaining probe locations. The curves obtained are similar and are not shown here.

In Table 2, the experimentally determined resonant frequencies are compared to the corresponding theoretical ones. In Table 2, the experimentally determined f_m were found as the average of the resonant frequencies (frequencies at which $G_{1,1}$ attains its maxima) for the four probe locations. It was mentioned above that the resonant frequencies differ from the average resonant frequencies at most in the last significant digit in Table 2. The theoretical resonant frequencies are found from the formulas in [1]. It is seen that the relative difference between theoretical and experimental f_m is very small, of the order of a fraction of a percent. The difference in general becomes smaller at the last resonances and the theoretical resonant frequencies are always smaller than those determined experimentally.

5. Conclusion

A swept-frequency measurement of the driving-point admittance of a circular array has been described. The single driven element is the extension of the inner conductor of a coaxial transmission line. The remaining 89 elements are brass monopoles fastened onto the aluminum ground plane. Narrow resonances are observed. The relative error between the measured resonant frequencies and those predicted in previous theoretical studies is very small.

driven element   parasitic elements

aluminum ground plane

voltage probe (to vector voltmeter)

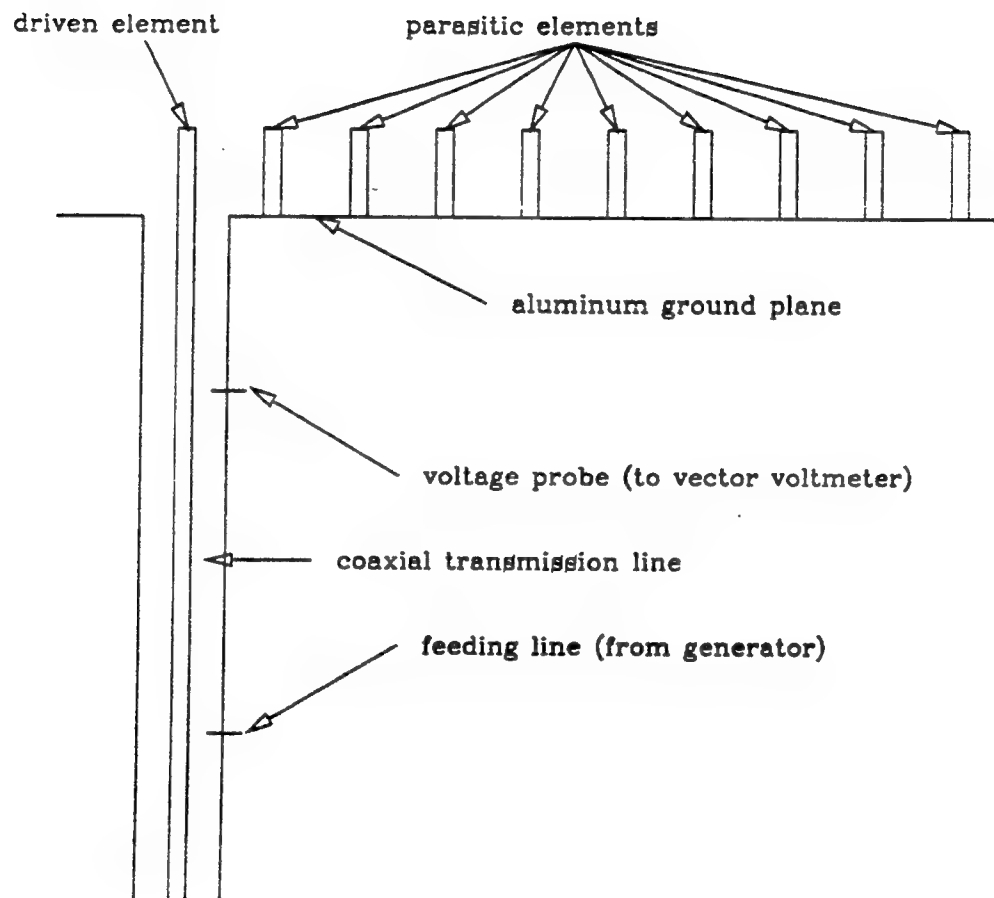coaxial transmission line

feeding line (from generator)

Figure 1a: General layout of the experiment; side view.
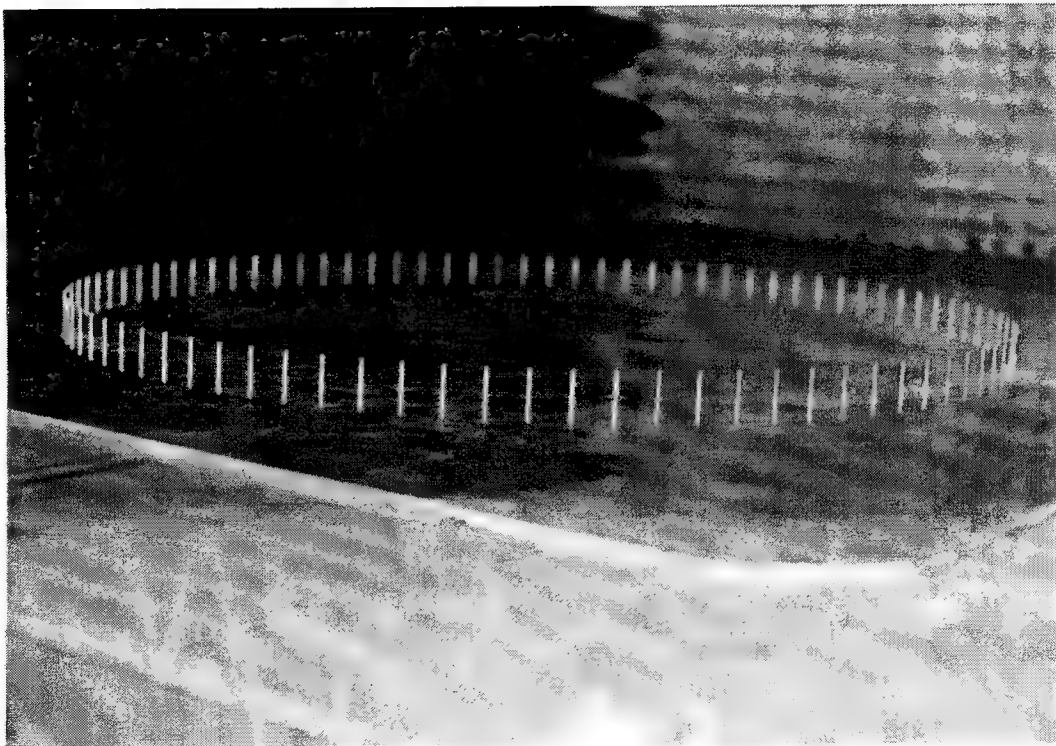
Figure 1b: General layout of the experiment; top view of the circular array.

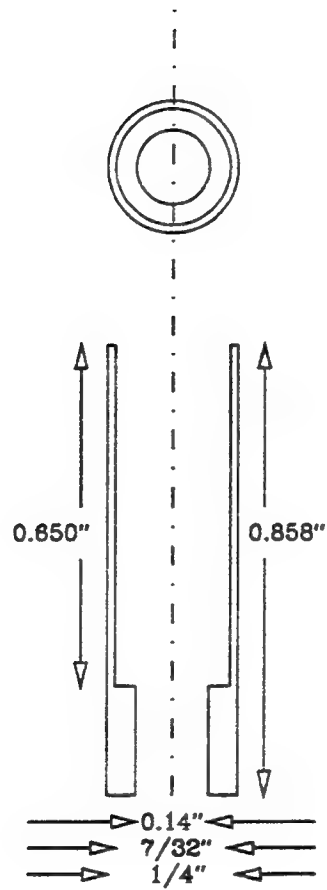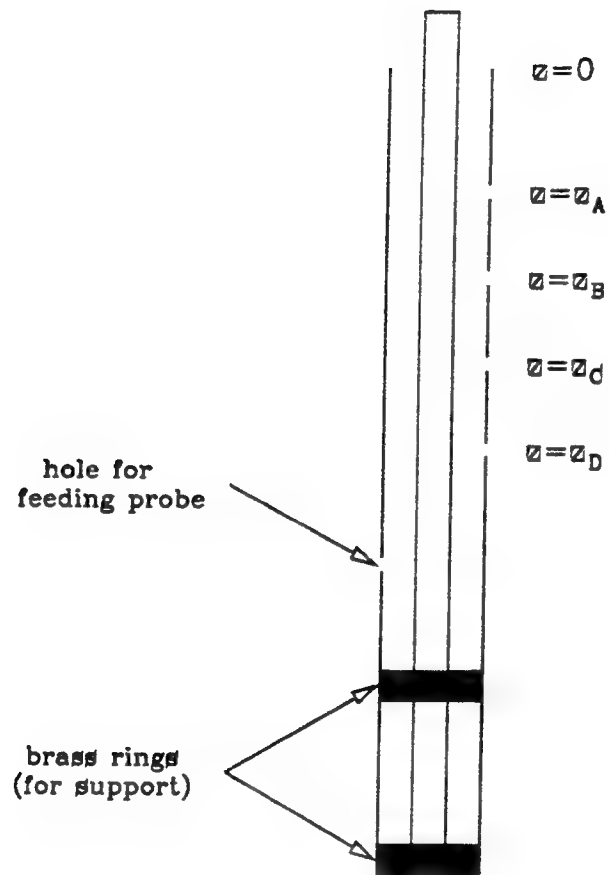Figure 2: Parasitic element.

z = 0

z = z_A

z = z_B

z = z_d

z = z_D

hole for
feeding probe

brass rings
(for support)

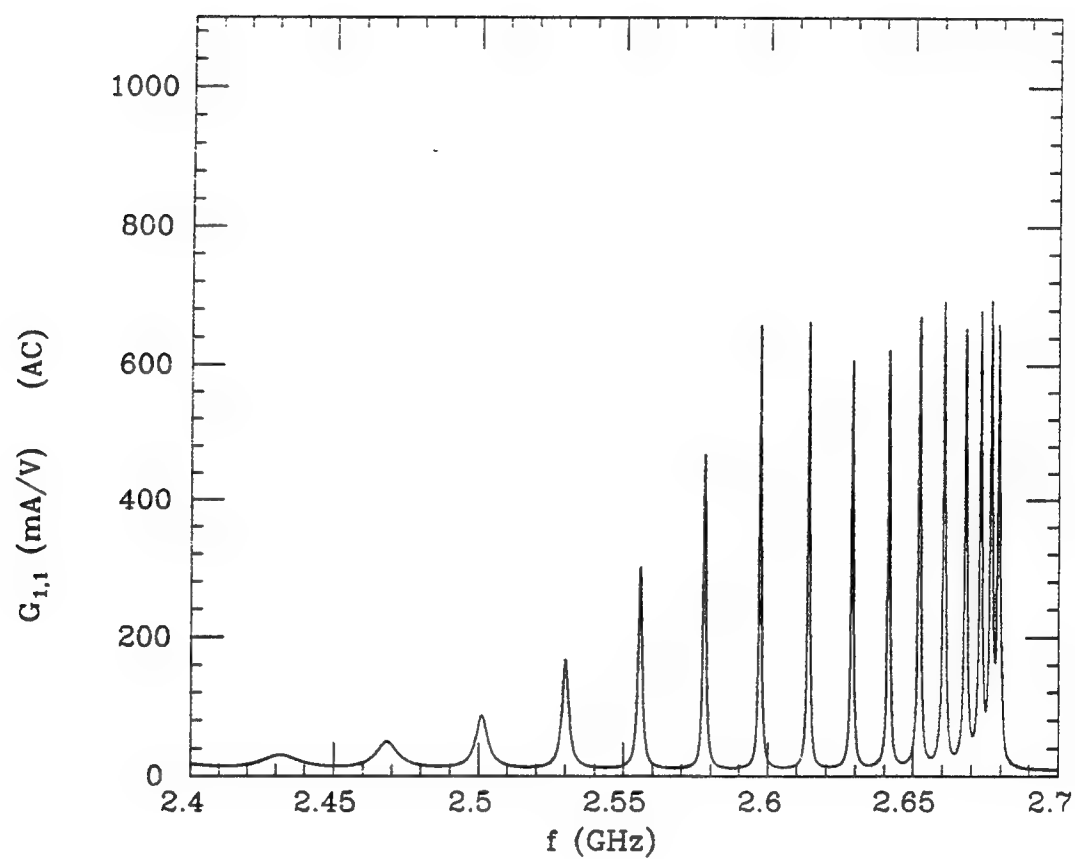Figure 3: Coaxial feeding transmission line–driven element.

Figure 4: Experimentally determined values of driving-point conductance $G_{1,1}(f)$ as a function of frequency for probe locations AC.
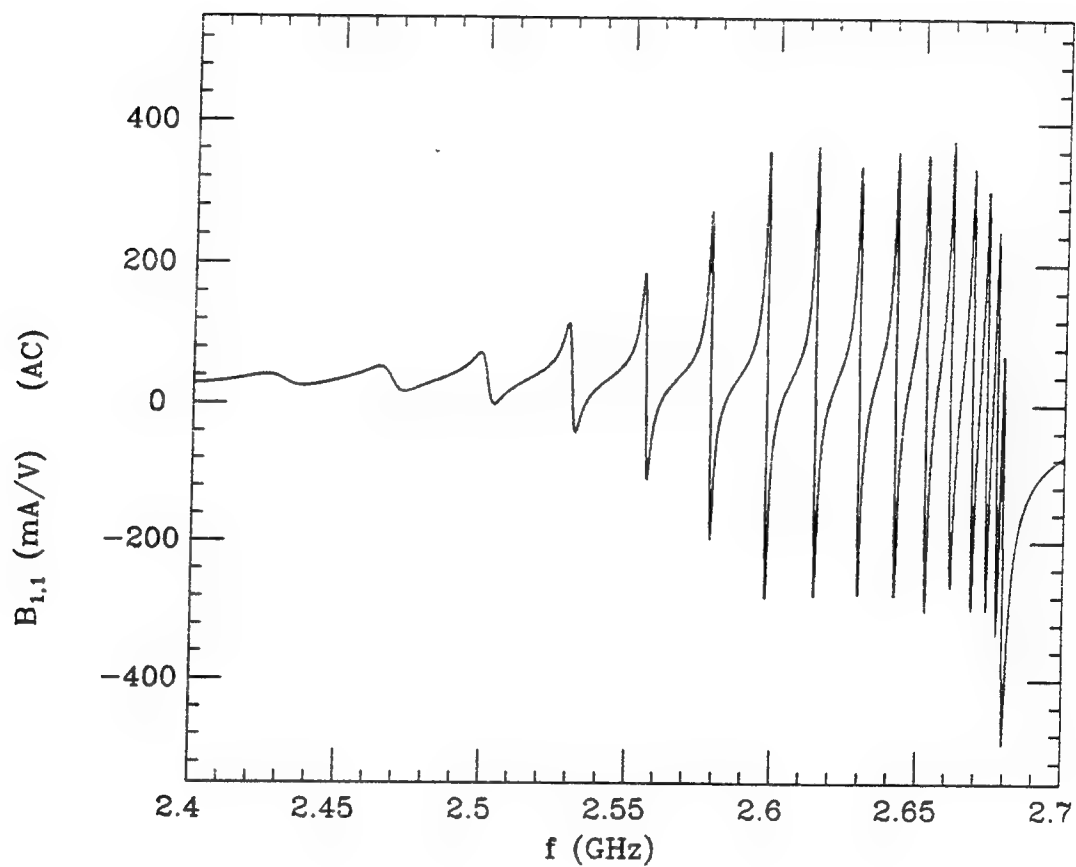
Figure 5: Experimentally determined values of driving-point susceptance $B_{1,1}(f)$ as a function of frequency for probe locations AC.

6. Acknowledgements

The author wishes to thank Professor R. W. P. King, Professor T. T. Wu, Dr. John Myers, Michael Brown, Hao-Ming Shen, and Dr. Richard Mack for valuable help and useful discussions.

7. References

[1] G. Fikioris, R. W. P. King and T. T. Wu, "A Novel Resonant Circular Array: Improved Analysis", in *Progress in Electromagnetics Research, vol. 8*, J. A. Kong, ed., pp. 1–30, New York: Elsevier, 1994.

[2] T. T. Wu, "Fermi Pseudopotentials and Resonances in Arrays", in *Resonances–Models and Phenomena: Proceedings, Bielefeld 1984*, S. Albeverio, L. S. Ferreira, L. Streit, eds., pp. 293–306, Berlin: Springer-Verlag, 1984.

[3] A. Grossmann and T. T. Wu, "A Class of Potentials with Extremely Narrow Resonances", *Chin. J. Phys.*, vol. 25, pp. 129–139, 1987.

[4] R. W. P. King, "Supergain Antennas and the Yagi and Circular Arrays", *IEEE Trans. Ant. Propagat.*, vol. AP–37, pp. 178 –186, Feb. 1989.

[5] R. W. P. King, R. B. Mack, and S. S. Sandler, *Arrays of Cylindrical Dipoles*, Ch. IV, New York: Cambridge University Press, 1968.

[6] G. Fikioris, R. W. P. King, and T. T. Wu, "The Resonant Circular Array of Electrically Short Elements", *J. Appl. Phys.*, vol. 68, pp. 431 –439, July 1990.

[7] H.-M. Shen, "Experimental Study of the Resonance of a Circular Array", presented at the *SPIE OE/LASE'91 Meeting*, Conference 1407 on Intense Microwave and Particle Beams II, Los Angeles, CA, 1991.

[8] D. K. Freeman and T. T. Wu, "An Improved Kernel for Arrays of Cylindrical Dipoles", presented at the *1991 IEEE/AP-S Symposium*, London, Ontario, June, 1991.

[9] D. K. Freeman, "Extremely Narrow Resonances in Closed-Loop Arrays of Quantum Mechanical and Electromagnetic Interactions", Ph.D thesis, Harvard University, 1992.

[10] G. Fikioris, "Resonant Arrays of Cylindrical Dipoles: Theory and Experiment", Ph.D thesis, Harvard University, 1993.

Array Failure Correction with a Digitally Beamformed Array

R.J. Mailloux
Electromagnetics and Reliability Directorate
Rome Laboratory

Abstract

A technique is presented for replacing the signals from failed elements in a digitally beamformed receive array. The technique is shown to be useful in a multiple signal environment and does not impose special restrictions on the correlation properties of the recieved signals. Experimental data confirm the operation of this algorithm for the case of one signal and one interfering source. The technique is shown to be accurate when the incident signal locations are not precisely known, and even when the signals are distributed over a broad angular region, or are only known to be located within a broad angular region.

1.0 INTRODUCTION

Failed elements can seriously degrade the pattern of low sidelobe arrays. A single failure in an N element uniformly illuminated array gives rise to a residual sidelobe level of $(1/N)^2$ below the main beam, corresonding to -40 dB for a 100 element array. Multiple failures can produce far greater levels. Recently there has been an increased interest in developing methods to improve the patterns of arrays in the presence of failed or partially failed elements by re-optimizing the weights applied to the remaining array elements[1,2].

This paper presents a new method for improving the receive radiation pattern of an array with failed elements. For the purposes of this work the elements will be assumed to be completely failed (and have zero signal output). This is the most interesting and useful case, because known errors in amplifier gain or phase shifter precision can be corrected by a simple complex multiply operation.The method can be applied in the presence of a multiplicity of external noise or jamming signals, and involves only linear processing that must be performed at a rate compatible with the information bandwidth . In its present form the technique assumes that all array active element patterns are

the same, and so the method is appropriate for large arrays or arrays without significant edge effects.

In previous papers [3,4] an algorithm was presented that used the signals from a small group of "good" elements to produce signals that could be substituted to replace the signals from failed elements in a large periodic array. The algorithm makes use of the fact that in a relatively large periodic array the signals recieved from any one far field source are periodic across the array and related by a complex constant. With a knowledge of that "phase constant" one can reproduce the signal that would have eminated from a failed element, and replace it by a signal that is constructed from the signals received at the other elements.

If a single wave is incident on the array, the process of replacement is trivial, but the special feature of the algorithm is its ability to separate out a number of signals , whether correlated or uncorrelated, that may be incident on the array from different directions.

The previous work [3,4] presented the algorithm, and examined the issue of how this correction altered array signal to noise, but otherwise assumed that the signals were all monotonic and that the phase constants were known exactly. In this case, and not considering receiver noise correlation properties, the correction is accomplished exactly, and for every signal processed in this way the effective array pattern is exactly that of the designed antenna pattern, with low sidelobes or nulls as designed. The utility of the method, as presented, was to discriminate against specific, locallized sources of interference. For this application it can improve system signal to noise ratio (S/N). The method was shown to be not useful if receiver noise is the dominant noise source.

An important problem of array correction however is to correct for failed elements in the presence of multiple received plane wave sources with frequencies and incident directions that are not precisely known, or may be only known to be present within a fairly broad angular region. For example, it would be

334

very important to correct the array response due to the presence of angle-distributed signals that result from terrain or vehicle scatter, which might occur as signal clutter. Maintenance of sidelobes against such incident signals would provide a major advantage. This paper addresses these important topics, and demonstrates how the technique can be applied for more general cases. The paper also investigates the response of the algorithm to unknown signals, by which we mean signals that are received by the array but whose presence was somehow overlooked, and so were not accounted for specifically in the algorithm.

2.0 BASIC ALGORITHM AND EXPERIMENTAL RESULTS:

If a single wave "p" is incident on an equispaced linear array antenna (Fig.1) from the angle θ_p ,with direction cosine u_p (for $u_p = \sin\theta_p$) , the signal at the n'th receiver port is:

$$e_n = n_n + F_n \tag{1}$$

where n_n is the receiver noise and F_n is the signal received by the n'th element. The signal F_n is

$$F_n = A_p X_p e^{jn\Delta_p} \tag{2}$$

where A_p is a complex constant (and includes the element pattern transfer function), and Δ_p and X_p are given by:

$$\Delta_p = 2\pi u_p \frac{d_x}{\lambda_p} \tag{3}$$

and

$$X_p = e^{j\omega_p t} \tag{4}$$

Although Δ_p is evidently a phase constant, it will sometiimes be referred to as a 'location' in reference to the angle parameter u_p

The digital beamformer weights each n'th port signal to form the array output signal E

$$E = \sum_n w_n e_n \tag{5}$$

335

where the w_n are the complex antenna weights , chosen to synthesize and steer a particular array pattern.

In the absence of noise, the signal at any m'th port is related to that at the n'th port, by the simple relation:

$$F_m = F_n e^{j(m-n)\Delta_p} \qquad (6)$$

If any m'th port were failed (zero output), and neglecting noise, one need only take the digital signal from the n'th port and perform the complex multiplication indicated above in order to reconstruct the complete (narrow band) time dependent signal at the m'th port. While this is true, it is shown in the Appendix that for one signal and uncorrelated receiver noise, this does not improve the received signal to noise ratio, and it does not beneficially alter the response of the antenna to any other wave from a different direction unless a similar procedure is carried out for that wave too. If however, the above procedure could be carried out for all incident signals, then the proper ratio of each of these signals to each other could be returned to that for an array without failures. An algorithm for performing this correction is presented here.

Given a number P of plane wave sources of radiation with complex amplitude A_p incident on the array , the output signal of the n'th antenna port is:

$$F_n = \sum_{p=1}^{P} A_p X_p e^{jn\Delta_p} \qquad (7)$$

and by inspection one can see that the required signal at some failed m'th element can be written in terms of the components of the signal at the n'th element as:

$$F_m = \sum_p [A_p X_p e^{jn\Delta_p}] e^{j(m-n)\Delta_p}. \qquad (8)$$

where the terms in the brackets are the received signals from the incident waves at the n'th terminal. One must separate these individual terms from

the combined received signal in order to use equation 8 for reconstruction at element m. This is accomplished by the following iterative scheme that will illustrate the procedure for solving for the bracketed contribution for $P > 1$.

Starting with a section of the array where it is presumed there are no failed elements, and assuming that one knows the angle of arrival and center RF frequency (or the RF phase progression Δ_p across the array) for each of the sources, then one can eliminate the signal due to one of the waves (say p=i) at any n'th element by combining the signal from the n'th element and part of the (n-1)st element. To carry out this procedure, one forms the following expression at each of P elements, starting at the n'th. At any k'th element:

$$F_k^1 = F_k^0 - e^{j\Delta_i} F_{k-1}^0 \tag{9}$$

for k=n, n-1, n-2....,n-(P-1).

In this expression, the superscript 1 indicates the result of the first iteration, and the superscript zero represents the received signals F_n. The results of the first iteration can be shown to be:

$$F_k^1 = \sum_p C_i(p) A_p X_p e^{jk\Delta_p} \tag{10}$$

for k= n, n-1,n-2,.......n-(P-2) where the constant $C_i(p)$ is given

$$C_i(p) = 1 - e^{-j(\Delta_p - \Delta_i)} = 2j e^{-\frac{j}{2}(\Delta_p - \Delta_i)} \sin[\frac{1}{2}(\Delta_p - \Delta_i)] \tag{11}$$

At any "in band" frequency, with d_x chosen to avoid grating lobes, $C_i(p)$ is only zero for the wave p=i.

Note that the equation 10 is similar to equation 7 except for the constant $C_i(p)$. Since $C_i(i)$ is zero, one signal contribution has been removed by the iteration. If only two signals (p=1 and i) were present, equation 10 would only

contain the signal for p=1, and one could directly evaluate $A_1 X_1 exp(jn\Delta_p)$. One could obtain the value of $A_i X_i exp(jn\Delta_1)$ directly by rewriting equation 9 to eliminate the wave with p=1 instead of the one with p=i. With the signals received by two adjacent "good" elements, one can therefore obtain $A_p X_p exp(jn\Delta_p)$ for each of the two waves, and one can use equation 9 to estimate and replace the signal at any failed m'th element.

If P signals are present, one would need to use P element ports, and perform P groups of P-1 iterations in order to isolate the contribution from one of the signals. Using S as the set of all P incident waves, the first iteration (Eq.10) results in P-1 signals F_k^1 at each of P-1 elements denoted by subscript k. The wave i is eliminated from these signals. Starting the second iteration with the P-1 signals F_k^1 and repeating the procedure results in a new set of signals F_k^2 at P-2 elements, with any two chosen signals eliminated.

Repeating this procedure, the q'th iteration, chosen to eliminate one signal "i' is indicated

$$F_k^q = F_k^{q-1} - e^{j\Delta_i} F_{k-1}^{q-1} \tag{12}$$

After q=P-1 iterations, in which all signals but one signal designated "s" are eliminated, one obtains at the n'th port:

$$F_n^q = \sum_p K_s(p) A_p X_p e^{jn\Delta_p} = K_s(s) A_s X_s e^{jn\Delta_s} \tag{13}$$

where $K_s(p)$ is given

$$K_s(p) = \prod_{i \in S; i \neq s} C_i(p) \tag{14}$$

and each $C_i(p)$ is given by equation 11. Note that the form of equation (13) is reduced because $K_s(p)$ is zero for $p \neq s$.

The iterative procedure is illustrated schematically in figure 2, and separates out one incident signal for each iteration. This procedure is essentially the same as the null forming tree of Davies [5] which forms nulls at a number of angles by means of a similar iteration. Here the groups of elements used

338

need not be the whole array,but must include at least P elements. Since $K_s(p)$ is zero for all of the waves except p=s, the contribution from waves at all up directions except u_s have been removed by the $K_s(p)$. Equation 13 therefore gives the remaining signal at the n'th element as:

$$[A_s X_s e^{jn\Delta_s}] = \frac{F_n^{P-1}}{K_s(s)} \tag{15}$$

which is one of the required terms (the one for p=s) in the brackets of equation 8. This procedure needs to be repeated(or processed in parallel) P - times in order to solve for all of the terms (for s=1 to P) needed in the various bracketed contributions in equation 8.

Thus, if one knows the phase constant Δ_p of all waves incident on the array, then one can separate out a factor which includes the amplitude, phase and the entire time dependence of each wave.

Figures 3 and 4 show the results of using this algorithm to process the experimental data received from an array of 32 elements, operating at several monotonic frequencies at C-band. In Figure 3 the solid curve is the measured array data with the array weights set to form a -40 dB Taylor taper. The dash-dotted curve is the same array with elements 1,3 and 10 failed. After using the correction algorithm, the dotted curve is obtained, and nearly coincides with the original solid curve. In this trivial case of a single source, the method almost perfectly reproduces the original pattern in any chosen look direction.

In Figure 4,two independent sources of equal amplitude are placed at 0 and -10 degrees. The three patterns again represent the original, failed and corrected array with the same failed elements as in Figure 3. The displayed pattern is the response of the array with a low sidelobe (-40dB) illumination, as that pattern is electronically scanned past the two sources. Notice that in this case the array without failures does not show a -40 dB response because two signals are sampled and displayed simultaneously, and so the sidelobes

339

can add together for any one data set. The corrected pattern shows that the correctionnearly reproduces the original array behavior.

3.0 FAILURE CORRECTION IN THE PRESENCE OF DISTRIBUTED SIGNALS OR INACCURATELY LOCATED SOURCES

The basic algorithm has been shown to operate well in a multi-signal environment, including real digital data with known source locations. That fact doesn't assure that the procedure is useful however, even though one can produce the signal suppression commensurate with a low sidelobe antenna pattern for these selected laboratory experiments. One could argue that with known incident signals one could (to a degree) re-synthesize pattern zeros to achieve similar suppression of unwanted signals. Moreover, with adaptive arrays, one could achieve S/N optimization in the presence of interference, whether there are failed elements or not.

This section discusses two cases that extend the utility of the technique, and so increase its value as an array control tool. In both of these cases,one assumes the angle of arrival (more specifically the phase constant Δ_p) is not precisely known, but may be approximately known. The algorithm cannot function exactly as above, because the individual incident waves cannot be completely eliminated at the end of each iterative sequence, Moreover,the received signal may be distributed over a range of angles, and the distributed signal may be correlated or uncorrelated over the angular region. In this case the problem is compounded because the algorithm is designed to operate with a finite number of incident waves with known locations , and for the distributed signal it may not be possible to expand the incident spectrum of waves in a finite series. Furthermore, one can only count on being able to find some small number of adjacent "good" elements at which to process the algorithm, The error introduced by using the algorithm for this generalized situation is described below.

Consider an array recieving a number of incident waves,that may include some discretization of a continuum of waves that occupies some given angular region.(It will later turn out that the ability to discretize the incident spectrum is an unnecessary restriction).

In this situation one operates the algorithm by assigning a number of angles as "sample points" so that a wave progression corresponds to that sample point "s" if $\Delta_p = \Delta_s$. For convenience let the set of all sample points be S, and let S' be the set of all remaining incident waves excluding those at sample points.This separation into two groups is not required, but performed for purposes of illustration since it contrasts to the procedure for all signal locations known, wherein the set S' is null.

Following the same iterative procedure, evaluating the expression (15) at each of the sample points, one obtains the resulting expression replacing (15)

$$[A_s X_s e^{jn\Delta_s} + \sum_{p \in S'} A_p X_p e^{jn\Delta_p} \frac{K_s(p)}{K_s(s)}] = \frac{F_n^{P_s-1}}{K_s(s)} \tag{16}$$

The difference between this equation and equation 15 is that none of the $K_s(p)$ are zero for any of the terms in $p \in S'$, so the summation at the right within the brackets is not zero, and one cannot solve for any $A_s X_s exp(jn\Delta_s)$ for known $\Delta_s (s \in S)$ in case it were present. However, one can evaluate the whole bracket for any particular s, and in fact for all values of $s \in S$. This corresponds exactly to the procedure, with the sumation terms representing an error, but also containing the signal information from all sources that do not lie exactly at sample points. The remaining issue is to determine the array pattern behavior at the sample points, and throughout the distributed region. .

Replacing (15) by (16) results in the expression below for the signal F_m that is reconstructed to replace the signal from any failed m'th element.

$$F_m \doteq \sum_{s \in S} [A_s X_s e^{jm\Delta_s} + \sum_{p \in S'} A_p X_p e^{jn\Delta_p} \frac{K_s(p)}{K_s(s)}] e^{j(m-n)\Delta_s} \qquad (17)$$

Although equation 17 obviously includes terms with the incorrect phase progression, the size of the error is not immediately apparent. The following sections illustrate the error and show that under certain circumstances it can be kept very small.

3.1 Correction for Two Incident Waves, One of which is Incorrectly Located

Assume a signal $A_1 X_1 exp(jn\Delta_{s1})$ has known location Δ_{s1} and another signal $A_p X_p exp(jn\Delta_p)$ is only known to be located near to Δ_{s2}. Beginning with two sample points Δ_{s1} and Δ_{s2}, and writing the equation for the reconstructed signal at the m'th element, one obtains:

$$F_m \doteq A_{s1} X_{s1} e^{jm\Delta_{s1}} + A_p X_p e^{jn\Delta_p} [\frac{K_{s1}(p)}{K_{s1}(s1)} e^{j(m-n)\Delta_{s1}} + \frac{K_{s2}(p)}{K_{s2}(s2)} e^{j(m-n)\Delta_{s2}}] \qquad (18)$$

For Δ_p very close to Δ_{s2} the second term in the bracket dominates, and the constant $K_{s2}(p)/K_{s2}(s2)$ is nearly unity and real, so the p'th signal is properly represented at the n'th element. That signal is transferred to the m'th element by the exponential term $exp(j(m-n)\Delta_{s2})$, which is not exact, since the correct number is $exp(j(m-n)\Delta_p)$ but is understood to be close. Since this introduces only a small phase error for (m-n) small, it is important that the base antenna element (n) is chosen close to the failed element(s). Thus in an array with multiple failures, it may be good practice to solve for the substitution signal several times, once near each group of failed elements, so as not to magnify the error caused by the incorrect phase constant. A second reason for solving the algorithm at several locations across an array with multiple failures is that if a single location were used the phase errors projected across the array would correlate, and the "subarray" formed by these signals

improperly reproduced at each failed element would cohere a beam at the angle of Δ_{s2}. It is much preferred that the small errors remain small, and be randomly distributed across the array.

For Δ_p very close to the sample point Δ_{s2}, so that

$$\Delta_p = \Delta_{s2} + \delta \tag{19}$$

where

$$\delta = 2\pi \frac{d_x}{\lambda} \delta u$$

then one can show that

$$F_m \doteq A_{s1} X_{s1} e^{jm\Delta_{s1}} + A_p X_p \left[-\frac{\delta}{2} \frac{e^{j[\Delta_p(n-\frac{1}{2})+\frac{1}{2}\Delta_{s1}+(m-n)\Delta_{s1}]}}{\sin(\frac{1}{2}(\Delta_{s2} - \Delta_{s1}))} + e^{j[-\frac{\delta}{2}-(m-n)\delta+m\Delta_p]} \right] \tag{20}$$

Of the three terms above, the first is the correct reproduction of the signal at Δ_{s1} at element m. The second has the wrong amplitude and phase, but is of the order of the error in angle estimation δu and so is small. The exponential dependence with m ($exp[jm\Delta_{s1}]$) results in a "subarray" formed by all "M" replaced signals at failed elements that will form a beam of amplitude $\frac{M}{N}\delta$, located at Δ_{s1}, assuming uform weighting relative to the main beam peak. Repeating the algorithm at several points in the array instead of a single point randomizes the phase errors, and reduces the normalized average sidelobe levels below the value of $\frac{M\delta^2}{N}$, again assuming uniform illumination. The third term in the above has $exp(jm\delta)$ multiplying the $exp(jm\Delta_p)$ dependence. Other terms in the exponent are of the order δ and so produce only small errors. Once again, the incorrect "m" dependence can be improved by exercising the algorithm at additional array locations, but this small error does not produce large sidelobes, only a subarray pointed at a slightly incorrect location if δu is small relative to the array normalized beamwidth parameter $\lambda/(N d_x)$.

3.2 Use of Multiple Sample Points to Reproduce an Angle Distributed Signal

Consider only one signal at $(\Delta_{s1} \leq \Delta_p \leq \Delta_{s2})$. The angular region between u_{s1} and u_{s2} is sampled at a number of equally spaced points 's' in the set S, starting at the end points. The fact that a single signal is chosen instead of a distribution offers no loss in generality, because the system is linear. This signal, translated to the m'th element, should be equal to $A_p X_p \exp(jm\Delta_p)$, but unless p corresponds exactly to one of the sample points 's', the reconstructed signal will not be exact. According to equation 17, the signal F, evaluated at some element m is given in terms of the signal $A_p X_p exp(jn\Delta_p)$ at element n by the expression:

$$F_m \doteq A_p X_p e^{jn\Delta_p} \sum_{s \in S} [\frac{K_s(p)}{K_s(s)} e^{j(m-n)\Delta_s}] \tag{21}$$

Seen in its simplest terms, this procedure is equivalent to the point matching expansion of an exponential function in terms of coefficients a_s multiplying a set of basis functions.

$$e^{j(m-n)\Delta_p} \doteq \sum_{s \in S} g(p,s) a_s \tag{22}$$

where the basis functions are:

$$g(p,s) = \frac{K_s(p)}{K_s(s)} \tag{23}$$

The "point matching" nature of the expansion is evident since g(p,s)=1 whenever p matches one of the sample points in the set S, say the point s=s1, and at each such point every other g(p,s) is zero for $s \neq s1$. The coefficients a_s are given as $exp[j(m-n)\Delta_s]$. The expansion is therefore exact for every sample

point p=s, and depending upon the number of sample points and the shape of the basis functions, can be quite accurate at other points within the angular region spanned by the sample points because of the interpolation or smoothing action performed by the basis functions. This will be discussed further in the next sections.

3.2.1 Generating a Correction Signal Using Two Sample Points

In order to make the procedure more clear, consider using two sample points , one at u_{s1} and one at u_{s2}, with $u_{s2} > u_{s1}$ and let the one unknown signal be at angle u_p between these limits.

Figure 5 shows the amplitude of these two basis functions for the case $u_{s2} - u_{s1}$ small. The specific data is for $u_{s2} = 0.5$ and $u_{s1} = 0.2$. Equation 9210gives the result for F_m approximated with two basis functions. The zeros in the basis functions assure that the approximation is exact at u_{s2} and u_{s1}, and the triangular shape to the basis functions aids in reconstructing the phase at Δ_p which is interpolated between the phases of the signals at the two sample points. For example, for a point u_p half way between u_{s1} and u_{s2} (at $u_p = (u_{s1} + u_{s2})/2$) one obtains, by expanding the basis functions in terms of the products $C_i(p)$:

$$F_m \doteq A_p X_p e^{jm\Delta_p} [2\frac{sin((\Delta_{s2} - \Delta_{s1})/4)}{sin((\Delta_{s2} - \Delta_{s1})/2)} cos(\frac{1}{2}(\Delta_{s2} - \Delta_{s1})(\frac{1}{2} - (m - n))] \quad (24)$$

Clearly the phase of this expression is exact at this value of Δ_p but there is amplitude modulation of the signal that makes this value of F_m less than it would have been for a perfect array. This modulation is the amplitude error,and depends upon the distance $(m - n)d_x$ between the base n'th element and the corrected element. This again is an argument for exercising the algo-

rithm at a multiciplicy of conveniently placed base elements.

For small $\Delta_{s2} - \Delta_{s1}$ this value of F_n is:

$$F_m \doteq A_p X_p e^{jm\Delta_p} \cos[\frac{1}{2}(\Delta_{s2} - \Delta_{s1})(\frac{1}{2} - (m-n))] \tag{25}$$

If one were to require that the modulation be 0.9 of the correct value, so that the sidelobes were 20dB lower than that for a single completely failed element, this would lead to the condition (for m-n sufficiently large):

$$|\frac{(\Delta_{s2} - \Delta_{s1})(m-n)}{2}| < 0.45 \tag{26}$$

but for smaller values of this argument both amplitude and phase are very accurate over the whole range. This example points out the interpolating role played by the triangular shaped basis functions.

If M elements were reconstructed, all from one base element location, they would have sidelobe levels less than $0.01M/N$.

3.2.2 Results of Using Multiple Sample Points

As indicated in equations (21) and (22), the use of multiple sample points in the vicinity of where signals are expected, leads to the proper replacement of these signals at failed elements because the function $\exp[j(m-n)\Delta_p]$ can be expanded in a series of normalized basis functions multiplying the exponent $exp[j(m-n)\Delta_s]$. This kind of expansion is commmonly used with sinc functions which, like the normalized $K_s(p)/K_s(s)$ basis functions,are unity at their peak and zero at the peak of every other function in the set. The sinc functions however are real, have low sidelobes, and asymptotic behavior that decays like the inverse of the distance to the peak. The normalized $K_s(p)/K_s(s)$ basis functions are complex and do not decay as strongly. Figure 6 (A. and B) shows the magnitude of the bases functions for $P_s=3$ and 4 over a wide range

with $u_{s2} = 0.5$ and $u_{s1} = 0.2$. These curves emphasize the similarity between these functions and the sinc functions. Larger values of P_s exhibit higher sidelobes, but nevertheless still produce increasingly better approximations with increasing P_s.

Figure 7 shows the computed pattern of an array with no failures (solid) and with elements 15,16,17 and 18 failed. Clearly the sidelobe level for the failed case is intolerable for an array designed for low sidelobes. Figure 8 shows the local sidelobe level for the same array using 5,7,and 12 sample points over the region between 0 and 0.5 in u-space. The solid curve is the original array, and the dashed curve is the array using 12 sample points. This data makes it evident that one can obtain a near perfect match if enough sample points are chosen.

4.0 RESPONSE TO A SIGNAL AT AN UNKNOWN LOCATION

The method presented has been shown to give an accurate substitution for waves with known locations, or approximately known locations, or as in the last example, for waves within a known angular sector.

The discussion has not considered the system response to waves from unknown locations, and not near any of the sample points that might have been assigned to handle individual or angle distributed signals. This consideration is of importance because these unknown signals will be processed with the rest of the incident waves; and the algorithm provides no assurance that that replica of any unknown signal will have anything like the correct amplitude and phase.

The results given in the next sections indecate that unknown signals can sometimes be processed with significant errors, depending upon their location

relative to existing sample points, and the number and spacing between the sample points. The presence of an unknown signal does not (in theory) result in an error in processing any known signals. These issues will be highlighted in the following examples, which lead to several conclusions about the handling of unknown signals.

4.1 Response to Unknown Signal when Correcting a Single Known Signal

If it has been assumed that only one signal (at Δ_{s1}) is incident on the array, then the algorithm reduces to the trivial case, and the signal at element n is asumed to be $A_{s1}X_{s1}exp[jn\Delta_{s1}]$. However. a if a second signal $A_pX_pexp[jn\Delta_p]$ is also present, and the existance of this signal is not known and therefore not accounted for in the algorithm, then the substituted signal at element m is:

$$F_m = A_{s1}X_{s1}e^{jm\Delta_{s1}} + A_pX_pe^{jn\Delta_p}\frac{K_{s1}(p)}{K_{s1}(s)}e^{j(m-n)\Delta_{s1}} \tag{27}$$

The first term can be shown to be the correct reproduction of the known signal at element m, and the second is a reproduction of the unknown signal with phase error $(m-n)(\Delta s1 - \Delta_p)$, and with a small amplitude error.

Clearly the phase error is significant since $(\Delta_{s1} - \Delta_p)$ can be on the order of π, and so somewhere in the pattern there are errors introduced in the processing of the unwanted signals. The error in the substituted current is however of the order of the unknown signal's current $| A_p |$ and so for a single element failure the pattern error is of the order of $(1/N)^2$ relative to the beam peak. This is of the same order as the error incurred by not processing the failed elements, so little has been lost by the procedure, while at the same time the procedure significantly improves the behavior for any known signals.

4.2 Response to Unknown Signal when Correcting Two or More Known Signals

The new algorithm is used directly when two or more known signals are incident at the array, or when a number of sample points are used to define a region of controlled pattern level. In this case the response to a single signal $A_p X_p exp[jn\Delta_p]$, with Δ_p not corresponding to one of the sample points Δ_s , is given directly by equation 17 as:

$$F_m = [A_s X_s e^{jm\Delta_s} + A_p X_p e^{jn\Delta_p} \sum_{s \in S} \frac{K_s(p)}{K_s(s)}]e^{j(m-n)\Delta_s} \tag{28}$$

Here again the first term is the known signal, reconstructed at element m. The second term has the time dependence of the unknown signal, but multiplied by an incorrect phase term and the summation. This term is multiplied by the ratio $K_s(p)/K_s(s)$ which introduces catastrophic failure under certain circumstances.

The ratio $K_s(p)/K_s(s)$ has an upper bound that is important when the unknown signal is far from any region of sampled signals. This is:

$$max \mid \frac{K_s(p)}{K_s(s)} \mid < \mid \frac{1}{min[\prod_{i \in S \; ; \; i \neq s} C_s(i)]} \mid \tag{29}$$

If the sample points are closely grouped together, this is:

$$max \mid \frac{K_s(p)}{K_s(s)} \mid < \mid \frac{1}{\prod_{i \in S \; : \; i \neq s} \frac{1}{2}(\Delta_s - \Delta_i)} \mid \tag{30}$$

This multiplier can be huge, under conditions of many closely spaced sample points. The previous example of 12 sample points between u=0 and u=0.5

349

, for s=6 leads to the multiplier in excess of 6 x 10^{12}. Clearly, with such a huge multiplier, any unknown signal with angular location far from the sampled area, can cause numerical instabilities in the results, especially when dealing with real data.

One should point out that the result is still correct at each sample point, but the response to an unknown signal far from the cluster of sample points can be significantly in error.

This undesired response can be improved in several ways. Instead of requiring a near perfect match to the ideal sidelobe level throughout the sampled region, and incurring the need for many sample points over a specified region, it is prudent to choose a compromise number of sample points that will reduce the error should an unknown signal be present. Secondly, it is important to use a signal locating algorithm (like MUSIC) that will locate all the sources and eliminate the problem of a collimated signal being treated as an "unknown". Thirdly, one could readily build in a numerical test on the absolute value of the calculated correction signal, which should be on the order of those measured at the "good" elements. If an unreasonably large signal is computed, one can simply not substitute it as a correction, but instead tolerate the pattern distortion caused by the failures until the unknown signal angle can be estimated.

5 .0 CONCLUSION

This paper has presented a detailed discussion of a new method for substituting corrected signals to replace the signals from failed elements of an array in which the beamforming is performed digitally. This paper extended the technique to consider performance in several limiting cases, including the case wherein one signal is only approximately located. In this case the error consists of both amplitude and phase errors, both on the order of the error

350

in the initial angle estimation. This simple case did however point out the importance of performing the algorithm at several places across the array in order to avoid producing a correlated error that could deteriorate performance.

A second extension of the technique allowed consideration of signals within or distributed over an angular region. This was accomplished by processing the algorithm at a number of sample points within the region, in addition to individual sample points at any known discrete sources. This procedure yielded a highly accurate reproduction of the ideal pattern within the sampled region, but was shown to produce a severely distorted response to any "unknown" signal that was not explicitly treated within the algorithm.

APPENDIX: CORRECTION IN THE PRESENCE OF RECEIVER NOISE

Obviously the perfect correction of signals from all failed array ports would lead to a perfect array, even with multiple failures. Because this is counterintuitive, it is helpful to consider the limitations imposed by receiver noise.

For simplicity, assume that tht signal plus noise at the n'th slement is developed for a single incident signal with analog to digital converter quantization effects neglected. In this case the array signal to noise ratio is:

$$\frac{S}{N} = \frac{\overline{E_s^* E_s}}{\overline{E_N^* E_N}} = \frac{W^\dagger M_s W}{W^\dagger M_N W} \qquad (A1)$$

where the superscript $\dagger$ means conjugate transpose, and the overbar indicates a time average. M is the covariance matrix.

$$M = \overline{e^* e^T} \qquad (A2)$$

Assuming that the noise terms are not correlated with the incident waves,

351

and the incident waves are not correlated with each other, the terms in the noise covariance matrix M_N are

$$m_N(s,t) = \mid n_s^2 \mid \delta(s,t) + \overline{n_s^* n_t}\, [1 - \delta(s,t)] + \sum_{p \neq 1} \mid A_p \mid^2 e^{j\Delta_p(t-s)} \qquad (A3)$$

where $\delta(s,t)$ is the Kroneker delta function.

The expression still contains the noise contributions due to any interfering signals $(p \neq 1)$.

The signal power S includes only the contribution from the wave (p=1) and is

$$\mid A_1 \mid^2 \mid \sum w_t \mid^2 \qquad (A4)$$

For the array without errors and deterministic weights w_t, the S/N is given by the above ratio with the central term zero, because the noise due to the various receivers is assumed uncorrelated. With only a single incident signal A_1 (that is all undesired signals A_p are zero), and all noise contributions equal, the S/N reduces to

$$\frac{S}{N} = \frac{\mid A_1 \mid^2 \mid \sum w_t \mid^2}{\mid n_n \mid^2 \sum \mid w_t^2 \mid} = \frac{\mid A_1 \mid^2}{\mid n_n \mid^2} g_A \qquad (5)$$

where g_A is the array factor gain.

$$g_A = \frac{\mid \sum w_t \mid^2}{\sum \mid w_t^2 \mid} \qquad (A6)$$

352

Note that S/N here is the array factor gain times the S/N of an individual array element.

Te study the counterintuitive limit in which all but one receiver is assumed to have failed, and all other signals are reconstructed from data of a single receiver, the S/N is first cast in the more general form below which assumes a single external signal, and separate receivers.

$$\frac{S}{N} = \frac{|A_1|^2 |\sum w_t|^2}{\sum_s \sum_t [n_s^* n_s \delta(s,t) + (1 - \delta(s,t))\overline{n_s^* n_t} w_s^* w_t]} \tag{A7}$$

If it is assumed that the receiver noise is now completely correlated, as it would be if all signals were constructed from a single signal, the S/N reduces to

$$\frac{S}{N} = \frac{|A_1|^2}{|n_n^* n_n|} \tag{A8}$$

Here, not surprisingly, the S/N is the same as that of an individual element. This test case gives some indication of the limitations of the technique. The signals from a number of failed elements can be reconstructed, but with a generally decreasing S/N level as more elements fail. In the limit, when only one signal is real and all the others are simulated, S/N is the same as for a single element.

ACKNOWLEDGMENTS

REFERENCES

1. J. Simmers,H.L. Southall, and T. O'Donnell "Advances in Niural Beamforming in Proceedings of the 1993 Antenna Applications Symposium, Uni-

versity of Illinois, Sept 1993

2. T.J. Peters, "A Conjugate Gradient-Based Algorithm to Minimize the Side-lobe Level of Planar Arrays with Element Failures" IEEE Trans. AP-39, No 10, Oct.1991, pp 1497-1504

3. R.J. Mailloux "Phased Array Error Correction Scheme", Electronics Letters. 1 April 1993, Vol.29, No.7, pp 573-574

4. R.J. Mailloux "A Phased Array Error Correction Scheme" IEEE AP-S International Symposium, June 1993, pp 202-205

5. Davies, D.E.N., "Independent angular steering of each zero of the directional pattern of a linear array", IEEE Trans. AP-15, Mar 1967, pp.296,298
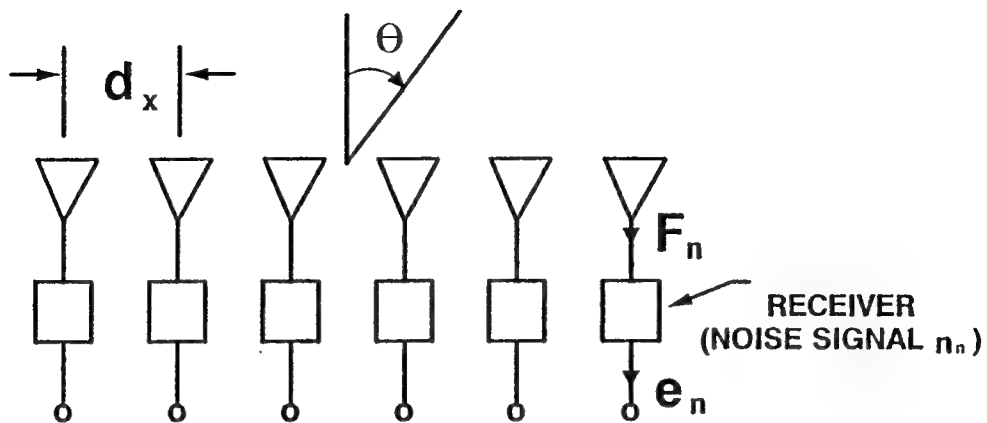
FIGURE 1: Linear Array Geometry

ARRAY CORRECTION ALGORITHM
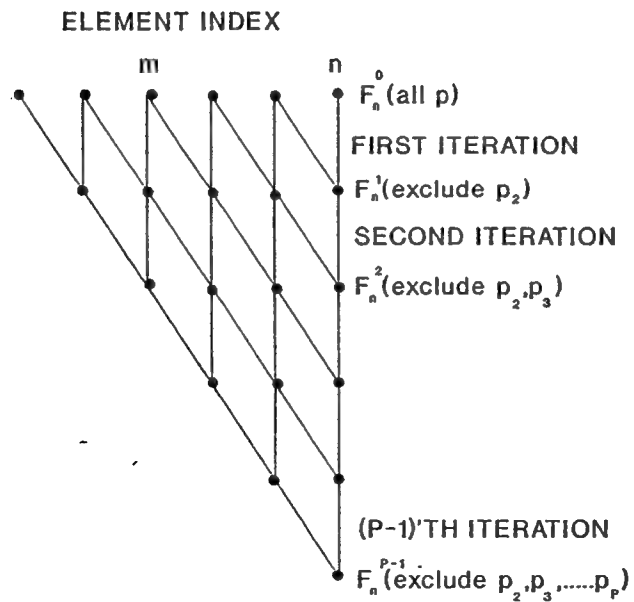
FIGURE 2: Array Correction Algorithm Flow Diagram
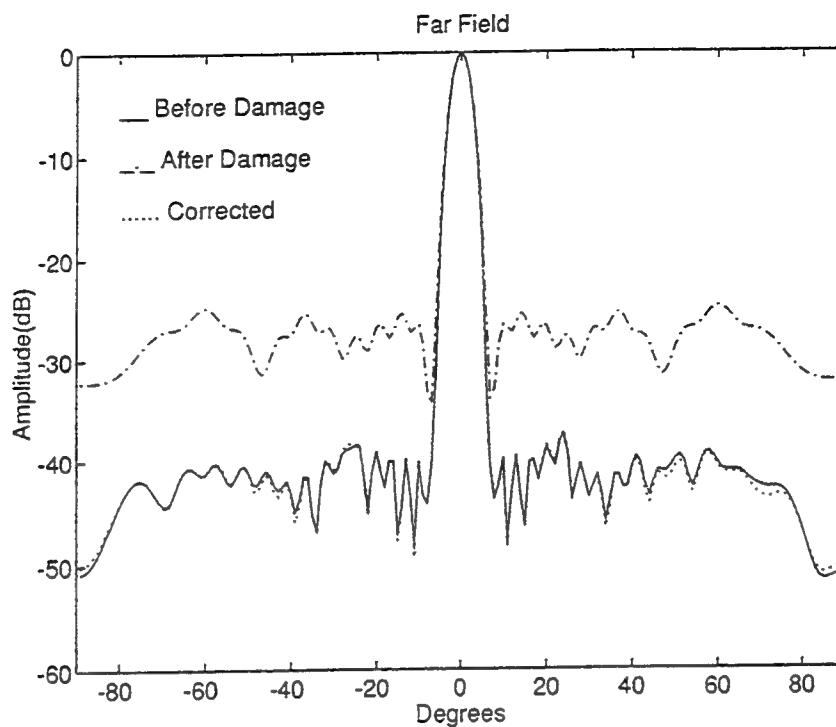
FIGURE 3: Experimental results for a single source at broadside
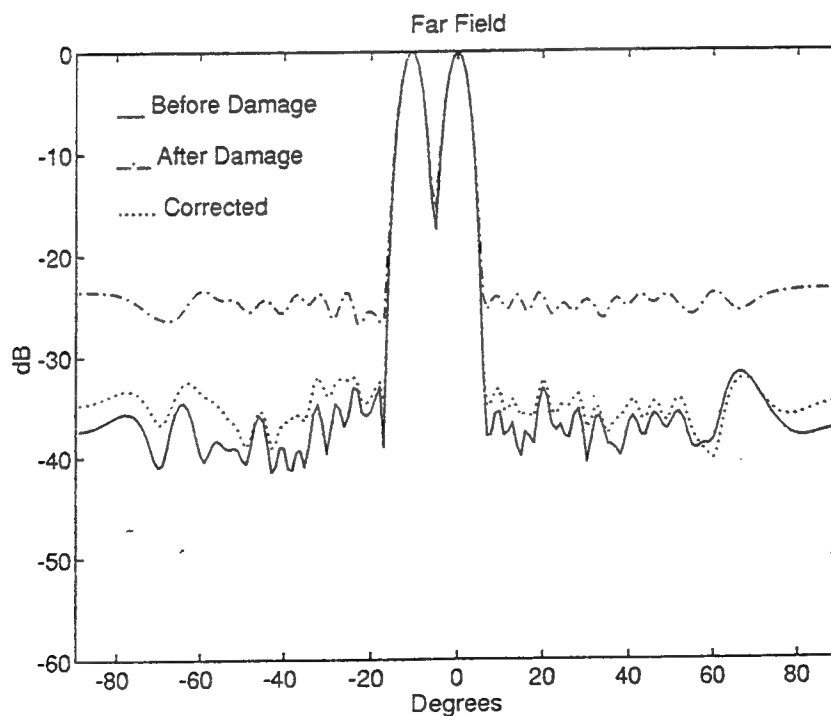
FIGURE 4: Experimental results for two uncorrelated sources, one at broadside and one at -10 degrees
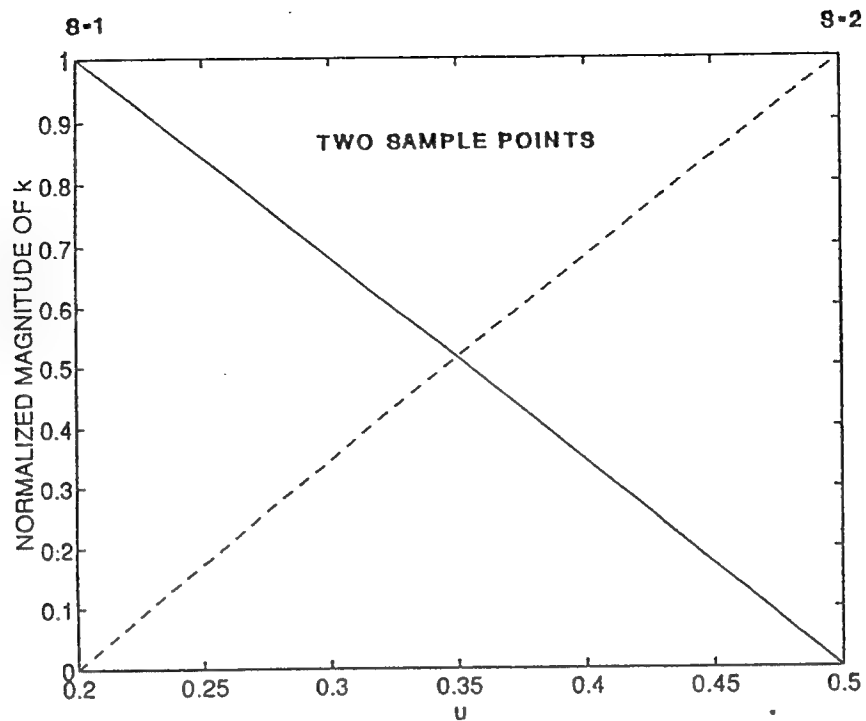
FIGURE 5: Basis functions (magnitude) for
representing a substituted signal
using two sample points
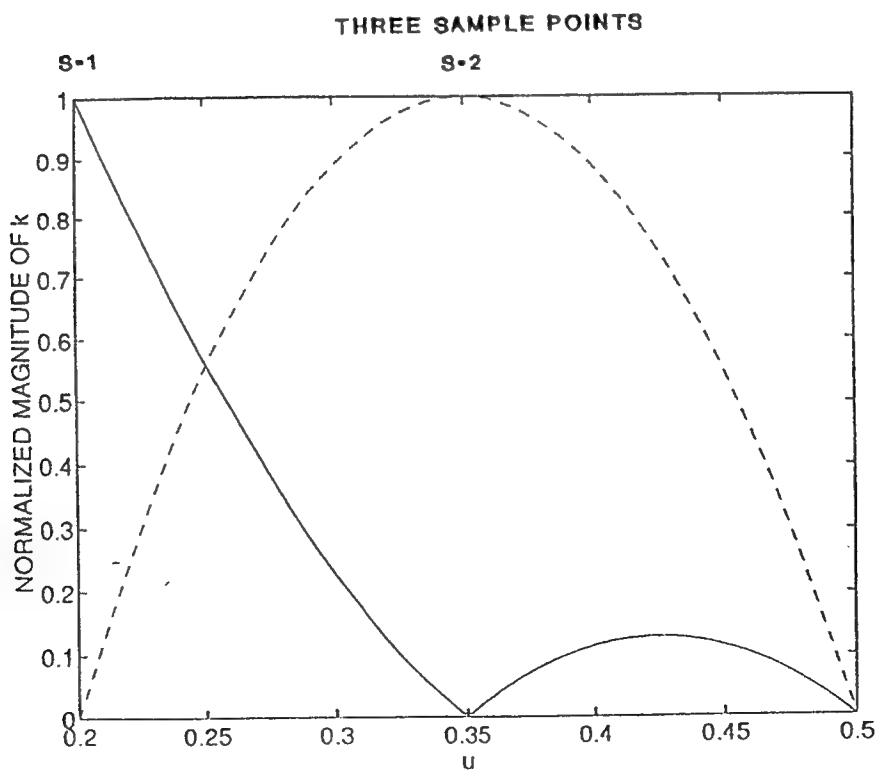
FIGURE 6a: Basis function (magnitude) for
representing a substituted signal.
Using three sample points (S=3 not shown)

357
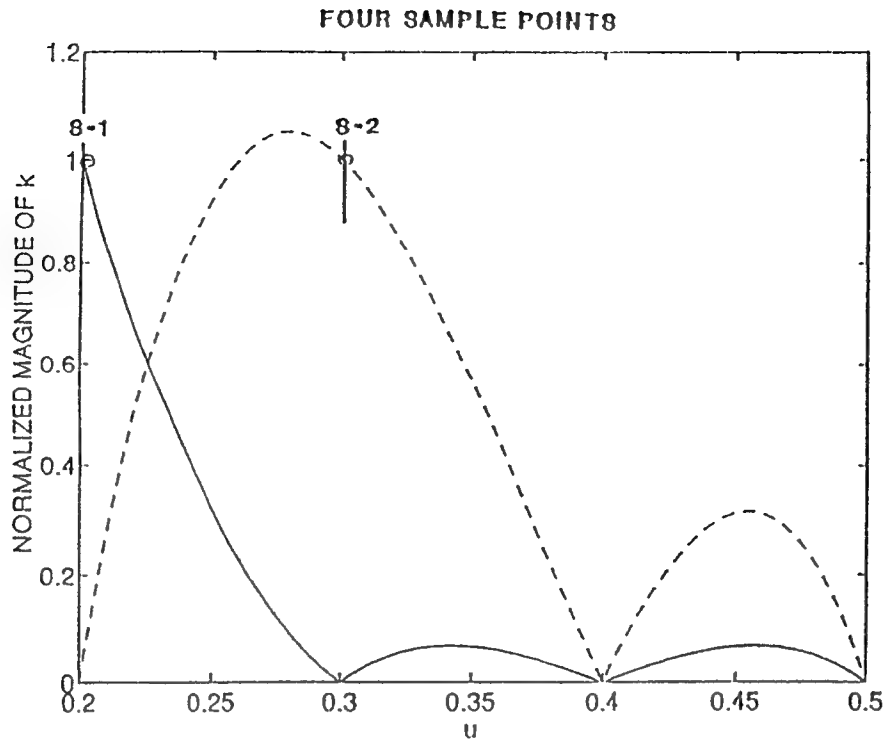
FIGURE 6b: Basis function (magnitude) for
representing a substituted signal.
Using four sample points (S=3 and 4
not shown)
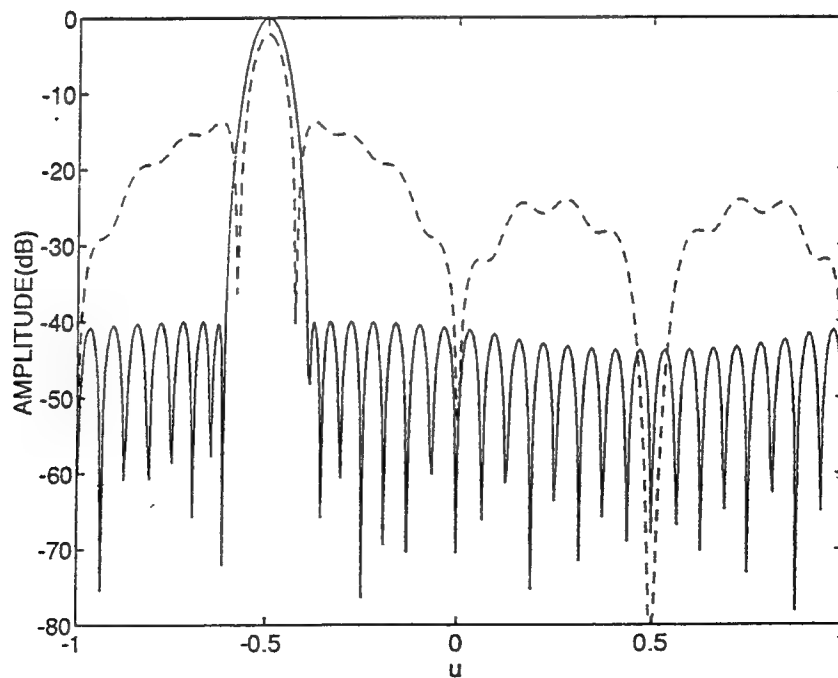
FIGURE 7: Computed pattern of ideal array
(solid) and array with 4 failed
elements (dashed)
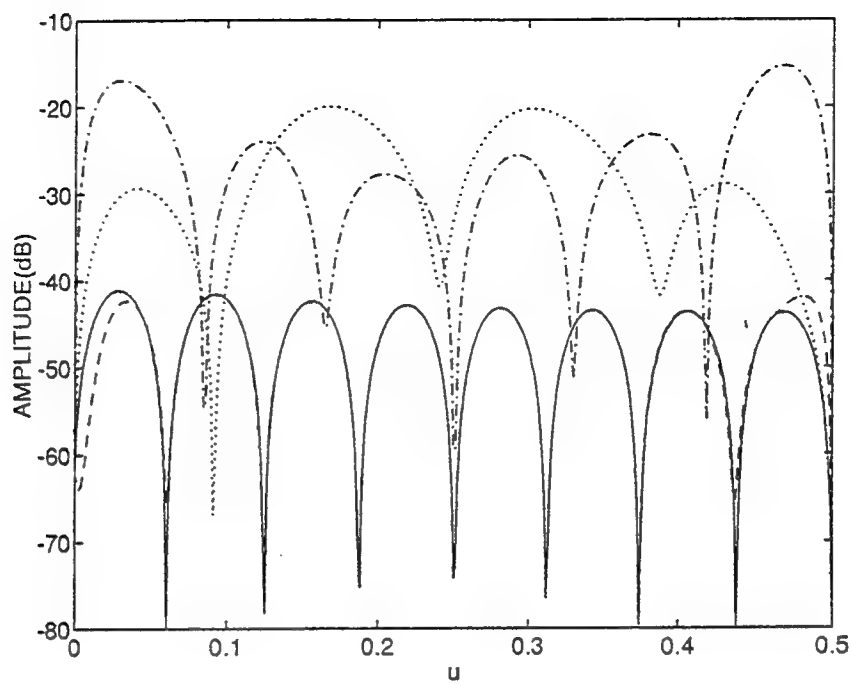
FIGURE 8: Pattern of corrected array in region of the sampled distributed signal:
———— Ideal array
······ Array with $P_S = 2$
—·—·— Array with $P_S = 7$
------- Array with $P_S = 12$

A DIGITALLY FOCUSED X-BAND IMAGING ARRAY

Stephen Frasier, James Mead, and Robert McIntosh
Microwave Remote Sensing Laboratory
University of Massachusetts
Amherst, MA 01003

Abstract

A unique phased array designed for rapid X-Band imaging of the ocean surface is described. The instrument, called the FOcused Phased Array Imaging Radar (FOPAIR), consists of a pulse-compression radar that employs a sequentially sampled (or synthetic) array of receiving elements. Range gating and software-based digital beamforming yield high-resolution, high-speed microwave imagery resolved in range and azimuth. A summary of the array design is presented. Due to its size, the array must be assembled and phase synchronized in the field. A technique developed by Attia and Steinberg that exploits the spatial correlation properties of the ocean surface has been found to work quite well for calibrating the array. Sample radar images illustrate FOPAIR's high-resolution ability.

1 Introduction

High resolution microwave imaging requires wide radiated bandwidths and physically large steerable apertures. Pulse compression techniques are often used to achieve fine range resolution while phased array techniques provide fine azimuthal resolution. The FOPAIR imaging radar uses these techniques to provide high resolution radar imagery of the moving ocean surface from a fixed platform such as a pier or tower. The array is sampled sufficiently fast to freeze the motion of the surface providing a nearly instantaneous complex radar image of the surface. These images are free of artifacts associated with surface motion during aperture synthesis frequently encountered in SAR. Rapid update of complex imagery at rates up to 180 Hz permits estimation of sea surface radial velocity and also permits transient surface features to be investigated.
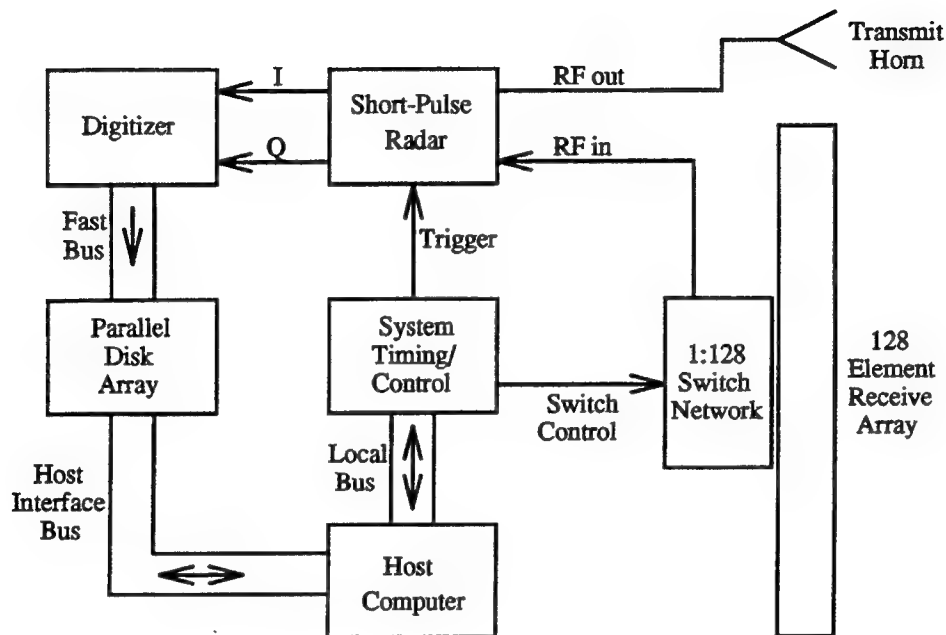
Figure 1: Block diagram of the FOPAIR radar system

Figure 1 shows a block diagram of the FOPAIR radar system. It consists of an X-Band pulse-compression radar that transmits through a pyramidal horn antenna illuminating the field-of-view to be imaged. Backscattered signals are received through a 128-element 6.8 meter long array controlled by a switching network that sequentially activates each element on a pulse-by-pulse basis. With each transmitted pulse, the signal received by the currently selected element is demodulated by the receiver, digitized by a high-throughput A/D converter, and stored to a fast parallel disk array. In this way, the radar receiver is time-shared by among all the array elements, and a burst of 128 transmitted pulses completes a scan of the array. At a typical PRF of 100 kHz the image capture time is 1.28 milliseconds.

Range resolution of 1.5 m is achieved by transmitting a 2 microsecond linear

FM chirp that is compressed to a 10 nanosecond pulse in the receiver. Azimuthal resolution of up to 0.25° is achieved by digitally focusing the array cross-range signals on a general purpose computer workstation. The digital beamforming includes a quadratic phase correction to focus the array into the near field and an FFT to produce an azimuthal scan of the array for the given range gate. Additional details on the radar system hardware can be found in [1].

2 The FOPAIR Array

The elements comprising the FOPAIR array are Linear Tapered Slot Antennas (LT-SAs) designed by the University of Massachusetts Antenna Laboratory. These elements are inexpensive broadband endfire antennas that are easily fabricated on printed circuit board [2]. The FOPAIR elements are fabricated on an inexpensive dielectric substrate (.031" Arlon 522 GT, $\epsilon_r = 2.55$) with a solder mask applied over the copper to protect the elements from possible corrosion. The LTSA's are 17 cm long with a 10° interior flare angle and are oriented to provide vertical polarization. Half power beamwidths are 18° and 24° in the E- and H-planes respectively, and typical directive gain is 14 dBi.

For ease of portability, the LTSA array elements are packaged into modules of sixteen elements each as illustrated in figure 2. Each antenna module also contains semi-rigid cabling and switching hardware necessary to address each of its elements. The cabling and switches are located on a mounting plate located directly behind the array elements. The module box and switch plates are fabricated of anodized aluminum with the exception of the front face consisting of a sandwich of fiberglass and low dielectric constant foam acting as a protective radome. The full array consists of eight such modules deployed end-to-end along a supporting frame fabricated from three 8-foot sections of aluminum radio tower.
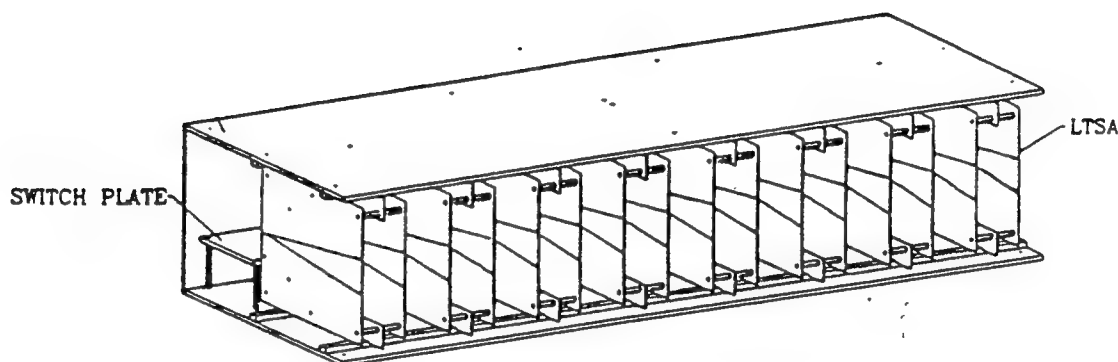
FOPAIR ANTENNA MODULE

Figure 2: Sample Array Module Cutout

The inter-element spacing in the array is 5.4 cm (1.8λ at 10 GHz) yielding an overall array length of 6.86 meters (229λ) and an approximate array factor beamwidth of $\lambda/L = 0.25°$. This spacing was chosen after a few design iterations to maximize azimuthal resolution subject to grating lobe rejection and mutual coupling effects. Grating lobes are located $\pm 32°$ in azimuth relative to the broadside beam direction limiting the array's visible space to $\pm 16°$. The grating lobes are rejected by the combined element patterns of the transmitting and receiving antenna elements that yield a gain pattern with a 3-dB beamwidth of 16°. The usable azimuthal footprint of the array is determined by the relative rejection of the grating lobe as the main beam is steered in azimuth. Grating lobe rejection of at least 20 dB is achieved when scanning is limited to within $\pm 12°$ off broadside.

The 1:128 switching network used to select individual antenna elements is implemented in three tiers as shown in figure 3. The first tier consists of sixteen single-
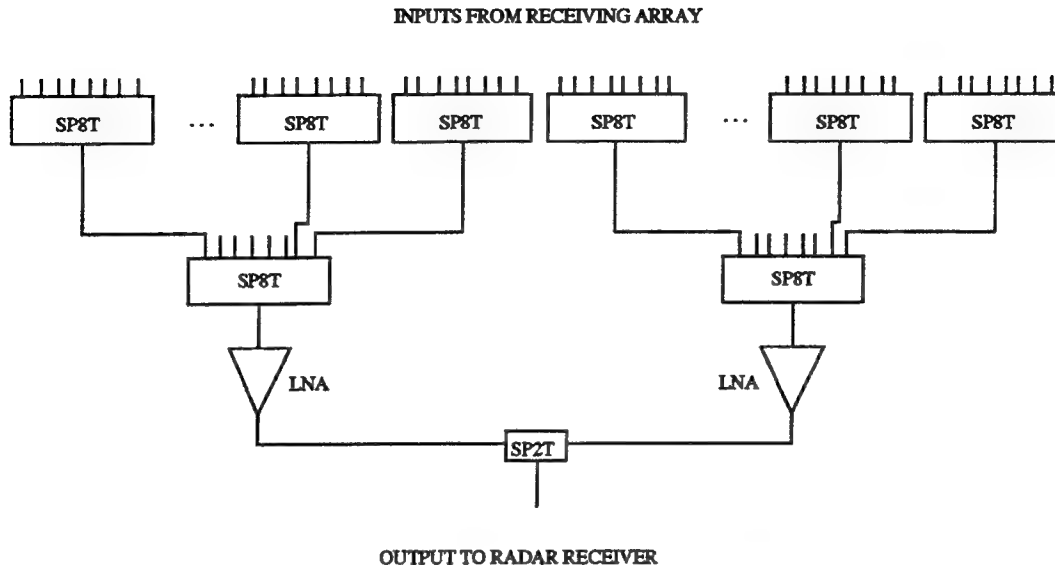
Figure 3: Receiving Array Switching Network

pole-eight-throw (SP8T) switches directly connected to the 128 array elements via short lengths of semi-rigid coaxial cable. The first tier switches are mounted on the aluminum plate located directly behind the antenna elements themselves. The outputs of these switches are connected via semi-rigid coaxial cable to a second tier consisting of two more SP8T switches that reside in two of the eight antenna modules. Each second tier switch services half of the array and feeds a low noise amplifier (LNA) to establish the noise figure for the receiver subsystem. Finally, two 12-foot flexible cables connect the two LNA outputs to a final SPDT switch in the radar receiver.

3 System Calibration

The performance of the focused array depends upon accurate phase and amplitude control of the array elements. Due to FOPAIR's modular array design, signals inci-

dent on the array elements travel through varying lengths of cable as they propagate through the switching network. This affects the relative time-of-arrival, phase, and amplitude of the signals received through the array elements, and these must be accounted for to focus the array properly.

Time-of-arrival differences are removed to an accuracy of 2.5 ns by the digitizer. This results in a maximum skew of ±20 cm in range between corresponding range-gates of adjacent array elements. The remaining phase and amplitude differences must be corrected through an array calibration procedure. Because the array is assembled on-site, field calibration is desirable. Repeated field calibrations also permit tracking of array performance.

3.1 Point-Target Calibration

The simplest technique to calibrate the array is to view a point target such as a corner reflector from a moderate distance. Given the range and relative azimuth of the corner reflector, the expected phase and amplitude variation across the array is known and can be used to derive a set of calibration coefficients, one for each element.

The point target calibration measurement is performed by illuminating a corner reflector placed nominally broadside at a range of 100-200 meters. This distance ensures that the beam reflected by the corner reflector uniformly illuminates the array, and that the point target itself lies entirely within one azimuthal resolution cell. Element-to-element phase and amplitude differences due to the array and switching network are embedded in the measured signal. They are corrected by dividing the measured complex signal by that expected response from a point target − a quadratically varying phase variation.

Figure 4 illustrates the results of such a field calibration. Here a corner reflector was deployed on a boat in calm water at range of 240 meters. The backscattered
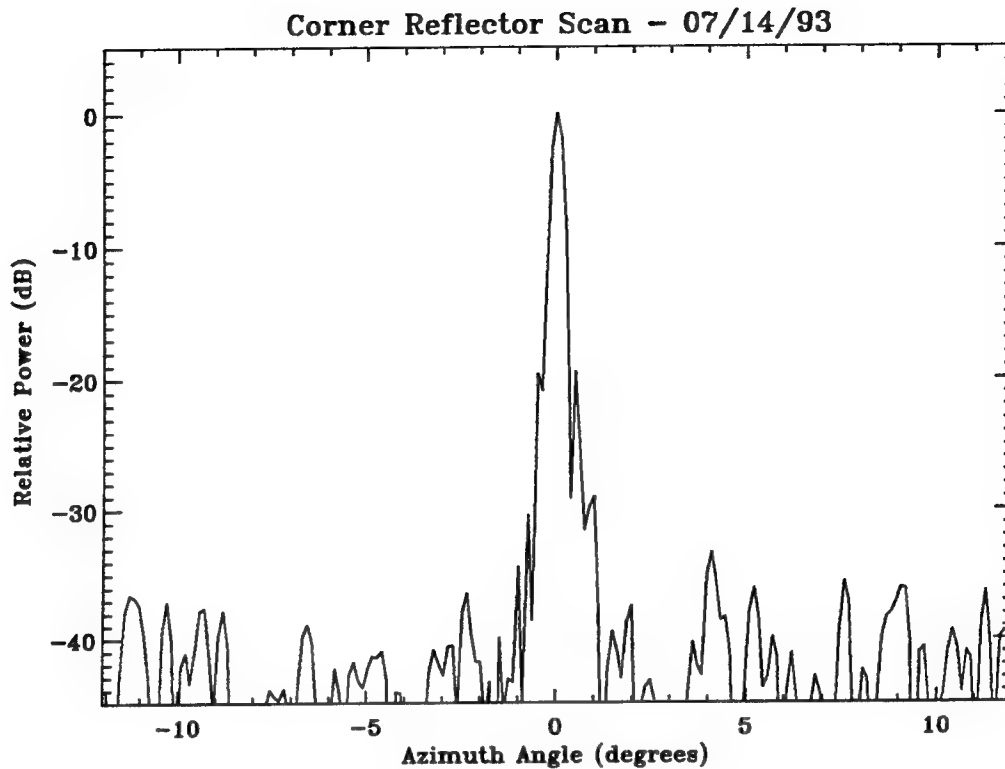
Figure 4: Scan of a Corner Reflector at 240 meters Range

signal was detected with an (unfocused) signal-to-noise ratio of about 25 dB, and a single realization of the backscattered signal (i.e. one array scan) was used to generate a set of calibration coefficients. A scan of the same corner reflector a few seconds later is shown in the figure. A cosine taper was applied to the field calibrated array response, and sidelobes away from the main beam are at least 33 dB below the main beam with average sidelobe levels somewhat lower.

The relative sidelobe levels of the array are determined by the quality of the field calibration measurement and upon the phase and amplitude error tolerances of the array elements. The quality of calibration depends upon how nearly the calibrating scene resembles a point source in free space. Although a corner reflector placed in the field of view is an adequate calibrating target, there always exists a finite signal-
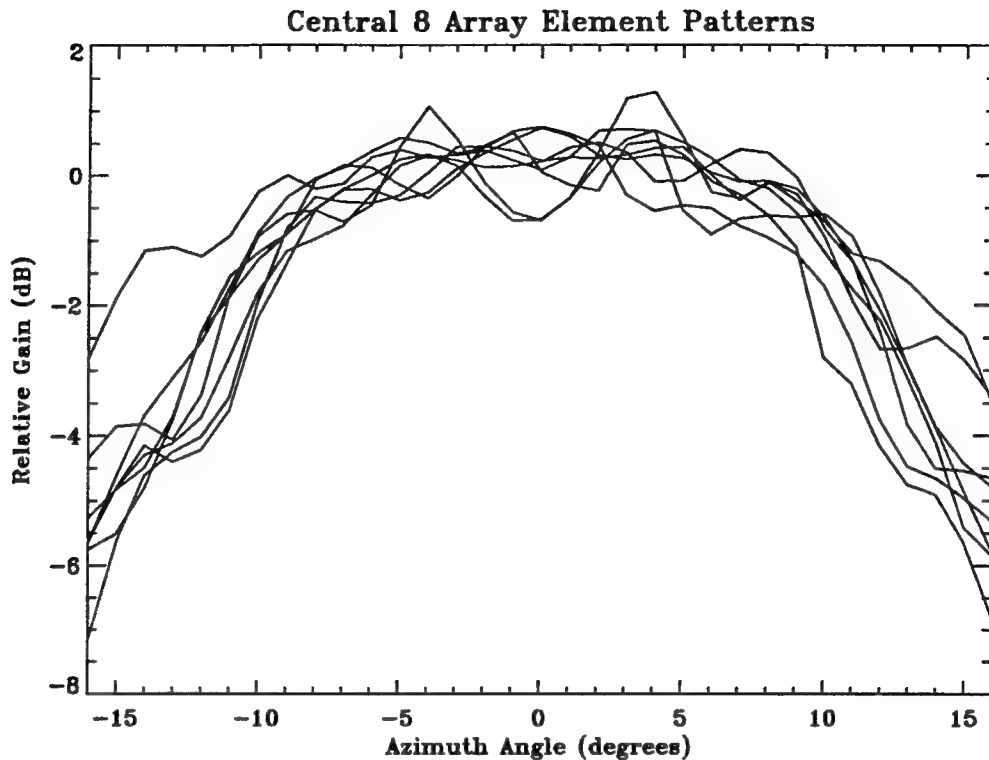
Figure 5: Measured gain patterns of the central 8 elements of a 16-element module

to-clutter ratio (SCR) or signal-to-noise ratio (SNR) that will give rise to random phase and amplitude errors in the received signal. Random errors such as these have the effect of limiting the achievable average sidelobe levels over the visible space of the array.

Additional errors due to variations in the fabrication of the individual LTSA elements, mutual coupling effects, and internal reflections are illustrated in figure 5. These antenna range measurements of the central 8 elements of a 16-element module show normalized amplitude variations of approximately 10%. Using an average sidelobe level estimate based upon small random phase and amplitude errors [3], average relative sidelobes should be expected between -35 to -40 dB for a 128 element array.

367

3.2 Clutter-Based Calibration Technique

In practice when FOPAIR is deployed for imaging the ocean surface, it is not always practical or possible to view a strong isolated point target. For example, when the calibrating target is deployed on a small boat, the motion of the boat due to the waves makes it difficult to point the reflector properly. Furthermore, other scatterers can cause interference corrupting the supposed spherical backscattered wave.

In the absense of a calibrating point source, the radar can also be calibrated using the statistical properties of the ocean surface backscatter. This technique, developed by Attia and Steinberg [4], exploits the decorrelation properties of a random clutter scene and the Fourier transform relationship between the scene and the field at the array aperture. The fundamental assumption of the technique is that the spatial autocorrelation function of the scene to be imaged is well described by a delta function, which occurs when an individual pixel is uncorrelated with it's neighbors. This is often assumed in microwave imaging of the ocean surface with pixel dimensions of the order of meters.

Under the assumption of a spatially uncorrelated clutter scene in the far field of the array, the spatial correlation function of the field at the array aperture is the Fourier transform of the azimuthal intensity distribution of the clutter scene, which is real and positive. If the scene is of uniform intensity, the observed intensity distribution describes the composite radiation pattern of the transmitter and receiver which is (hopefully) symmetric about the boresight direction. Since the Fourier transform of a real, and symmetric function is also real and symmetric, the spatial correlation function at the array should be real and positive for small spatial lags. That is, it should have zero phase. This spatial autocorrelation function can be computed explicitly by cross-correlating the signals from adjacent array elements and averaging over many range-gates and/or time samples to ensure good estimates.

Element-to-element variations in the phase of the calculated autocorrelation function are indicative of the relative phase differences between the elements due to the varying cable lengths, switches, etc. To account for near-field operation, it is sufficient to include a quadratic phase taper to the unfocused array responses prior to cross-correlating adjacent array elements.

The use of this technique was investigated during one set of ocean measurements and is compared to a point target calibration in figure 6. In this cases, only 64 elements of the array were used with a cosine taper applied. The first figure shows a corner reflector scan using the point target calibration, while the second figure shows the same corner reflector scan using a clutter-based calibration.

Both calibration methods show a well-formed beam in the direction of the corner reflector although the sidelobe levels appear to be generally worse for the clutter-based calibration. However, the sidelobe performance shown for the point-target calibration may be somewhat optimistic. That technique automatically optimizes for a point target at the relative azimuth of the calibrating target by equalizing the relative gains of the individual antenna elements for that specific azimuth angle. At other azimuth angles, however, the variations in the patterns indicated by figure 5 will contribute amplitude errors to degrade the sidelobe performance.

On the other hand, the clutter-based calibration technique uses the received power from the distributed backscatter over a much wider cone of azimuth angles. It is therefore not optimized for any specific look direction, but rather an average response across the field of view. In any case, the two methods show that relative sidelobe levels away from the main beam direction of better than -30 dB can be achieved through calibration in the field with or without a calibrating point source.

369
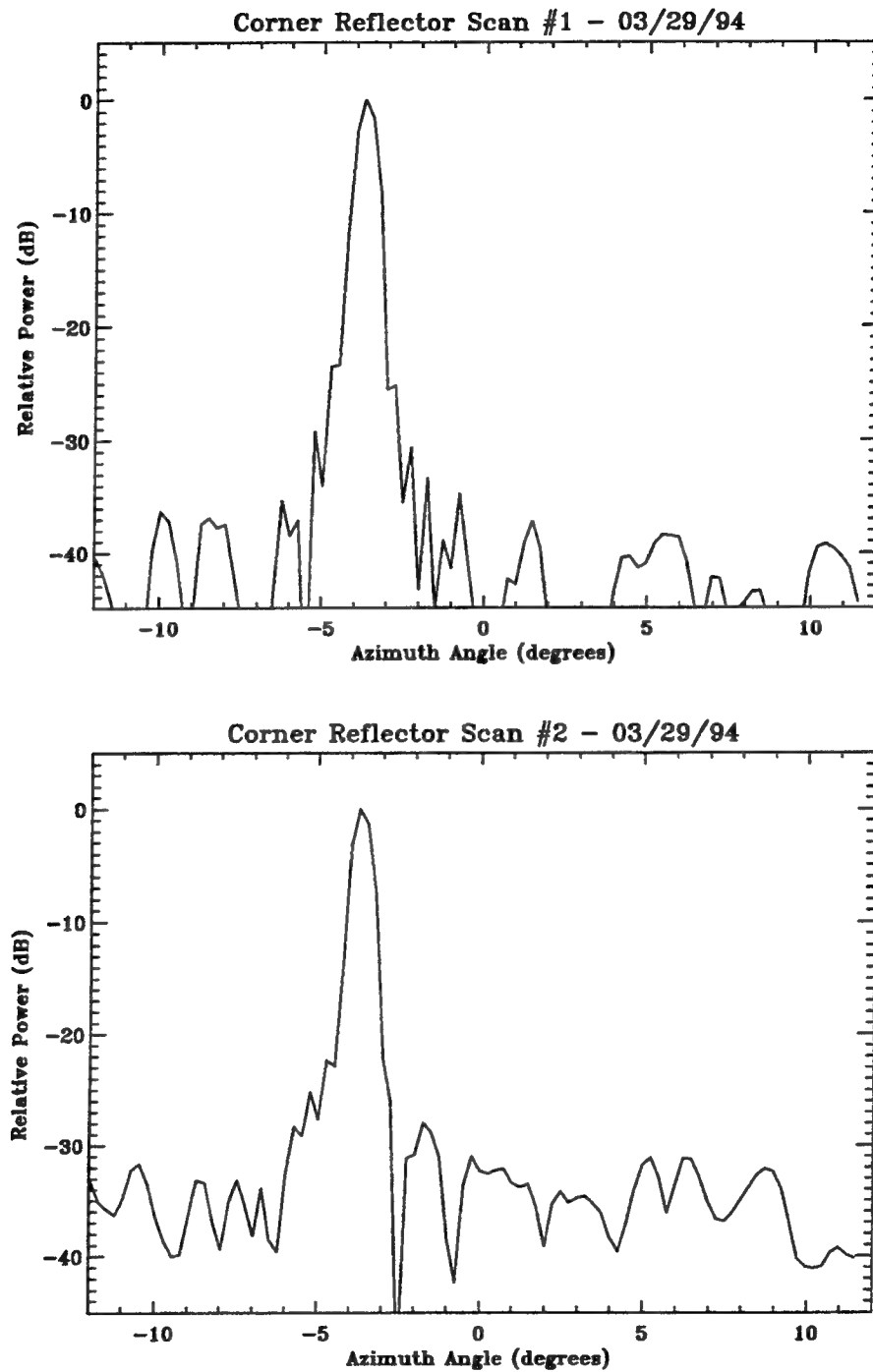
Figure 6: Scan of a corner reflector at 210 meters range using the point-target calibration procedure (top) and the clutter-based calibration procedure (bottom)
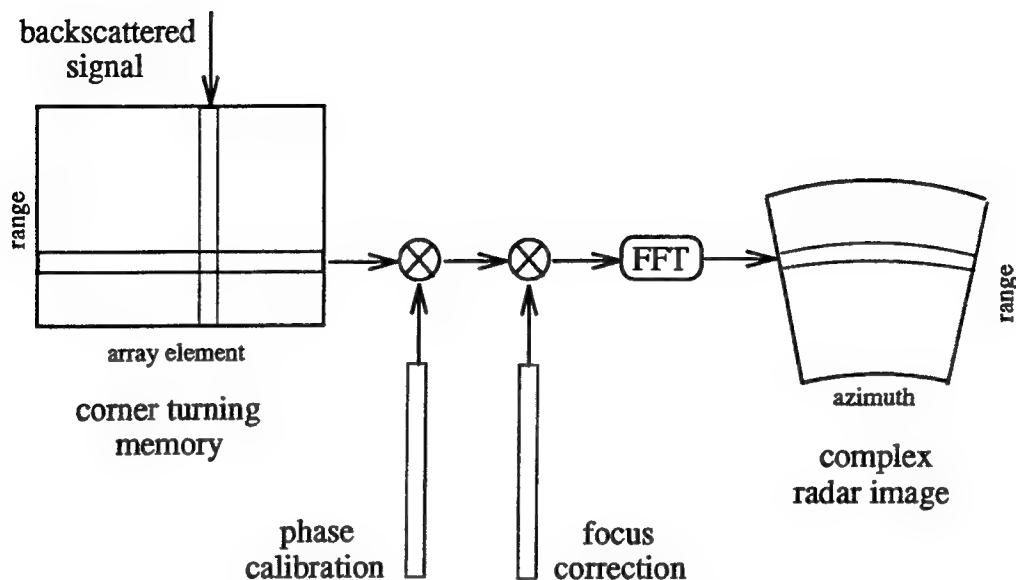
Figure 7: Flowchart Diagram of the image focusing algorithm

4 Beamforming

The processing algorithm used to form imagery is illustrated in figure 7. The detected in-phase and quadrature signals from the array are loaded into a 2-dimensional complex array indexed by antenna element and range gate. While this data array is filled element by element, the processing is done range-gate by range-gate. Thus, the data array serves as a corner turning memory. For each range gate, the samples are multiplied by the array's phase calibration and by an array taper that includes a quadratic phase variation to focus the array to the appropriate range. At this point, an FFT operation implements a scan of the array over the full visible space resulting in a single-look complex image of the scene.

To reduce the effects of fading (or image speckle), multiple scans of the array are incoherently averaged to provide despeckled, high-resolution images of backscattered power. An example of the resulting imagery is shown in figure 8. For this

371
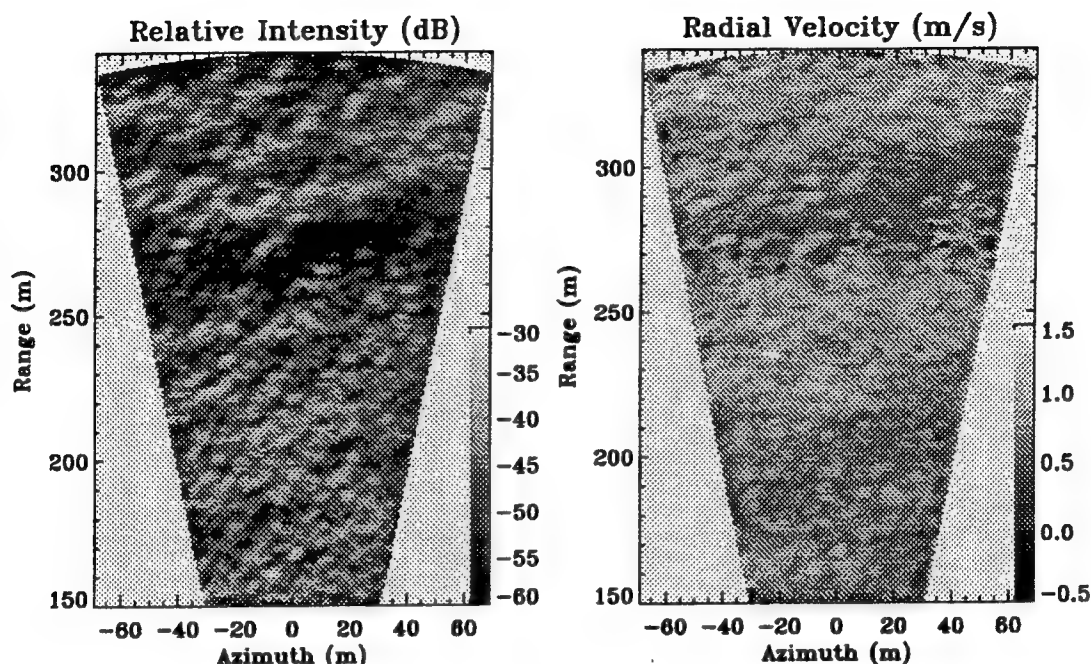
Figure 8: Sample FOPAIR intensity (left) and velocity (right) images obtained from a pier.

imagery, the radar was mounted on the end of a pier at a height of 12 m above mean water level and oriented directly seaward. The radar imaged a 24° wide sector between 150 m and 350 m range by using 64 elements of the array each sampling 128 range gates. The radar repeated scans of the surface at 180 frames/second, and 36 images were averaged together to produce the intensity image shown here. Thus this image represents a despeckled snapshot of the ocean surface obtained within 200 milliseconds.

The image is best interpreted as a plan-view of the ocean surface as illuminated from the side (in this case, from the bottom of the image). This image is corrected for the antenna element patterns and a cubic range rolloff, so the intensity is proportional to normalized radar cross section, σ°. Most evident in the image are the crests of diagonally propagating ocean waves. The dark band just beyond 250 m

range indicates the rear face of a low amplitude ocean swell that is propagating toward the shore (and the radar).

Because of the rapid image update, radial velocity can also be estimated at each pixel location. An efficient means of doing this is provided by the covariance or "pulse-pair" technique [5] that estimates the mean Doppler shift of a backscattered signal from the difference in the phase of the received signal over a fixed time lag. In our case, the phase of the cross correlation between two subsequent complex images is proportional to the surface velocity projected in the radar direction.

The right hand image in figure 8 shows the radial velocity corresponding to the intensity image. This image indicates directly the orbital velocities of the ocean surface waves. In particular, the larger scale ocean swell is quite apparent in this image. Such imagery can be Fourier transformed and converted to obtain a directional wave height spectrum.

Acknowledgements

References

[1] R.E. McIntosh, S.J. Frasier, and J.B. Mead. FOPAIR: A focused phased array imaging radar for ocean remote sensing. *IEEE Transaction on Geoscience and Remote Sensing*, 1994. in-press.

[2] K.S. Yngvesson, D.H. Schaubert, T.L. Korzeniowski, E.K. Kollberg, T. Thorn-gren, and J.F. Johansson. Endfire tapered slot antennas on dielectric substrates. *IEEE Transaction on Antennas and Propagation*, AP-33(12):1392–1400, 1985.

[3] M.I. Skolnik. *Introduction to Radar Systems*. McGraw-Hill, New York, 1980.

[4] E.H. Attia and B.D. Steinberg. Self-cohering large antenna arrays using the spatial correlation properties of radar clutter. *IEEE Transaction on Antennas and Propagation*, 37(1):30–38, 1989.

[5] K.S. Miller and M.M. Rochwarger. A covariance approach to spectral moment estimation. *IEEE Transactions on Information Theory*, 18(5):588–596, 1972.

Phased Array Antennas with Neural Network Signal Processing

Hugh Southall and Major Jeff Simmers
USAF Rome Laboratory
Electromagnetics and Reliability Directorate
RL/ERAS, 31 Grenier Street
Hanscom AFB, MA 01731

Terry O'Donnell*
ARCON Corporation
260 Bear Hill Road
Waltham, MA 02154

Abstract

We use a radial basis function artificial neural network to effectively perform the direction finding (DF) function when combined with an X-band, linear, phased array antenna. DF performance is presented for both simulated and experimental data and various signal to noise ratios (SNRs). The classical monopulse DF algorithm is used as a baseline for comparison. A new 32 element, microstrip patch array is used to obtain the experimental data. We compare performance of the full 32 element array and an eight element array using the center eight elements of the full array to obtain an estimate of the way DF performance scales with the number of elements.

1 Introduction

In this paper, we study DF performance of both a neural network beamformer and classical monopulse. In particular, we calculate rms angle error curves as a function of SNR for arrays of eight and 32 elements to see how both algorithms scale with the number of elements. This scaling is well known for monopulse; however, it has not been determined for a neural net beamformer. We use both simulated data (perfect antenna with added thermal noise) and experimental data to predict DF performance.

We have previously introduced two neural beamformers based on radial basis functions used for single source DF [1, 2, 3]. Our paper begins with an overview of a radial basis function neural beamformer. The classical monopulse algorithm is described next. We describe the experiment, present rms angle error curves, and make some concluding remarks.

2 The Neural Beamformer

The main components of the beamformer include preprocessing, an artificial neural network, and postprocessing. Preprocessing and postprocessing configure the network interfaces to perform a particular function, in our case direction finding. The neural network performs a function approximation role. Function approximation is described shortly when we discuss network training.

The choice of network architecture is crucial for successfully implementing a neural network in a practical application. We chose radial basis function networks because they can theoretically model any continuous function[4]. Single source DF techniques imply the existence of a continuous function between the space of input vectors, $\hat{X}$, and the space of possible angular directions to the source, $\hat{Y}$, i.e. $f : \hat{X} \rightarrow \hat{Y}$. Our empirical work has shown this mapping to remain relatively continuous despite many forms of antenna degradation and failure, and a variety of near field scattering conditions.

Much of the useful information for direction finding lies in the phase differences between elements as a result of a plane wave hitting the array aperture at angles other than broadside. During preprocessing we compute phase differences between consecutive pairs of elements and the network accepts these phase differences as components of an input vector. For the eight element array, input vectors therefore have seven components, the seven phase differences. Similarly, the 32 element array has 31-component input vectors.

Input vectors are mapped to an output space representing source azimuth angles off antenna broadside. We use 13 output nodes denoting azimuth angles from $-60°$ to $+60°$ in $10°$ increments. Postprocessing considers these 13 nodes as 13-component output vectors. Each output vector is associated with a single input vector. We present the network with known source azimuth angles (typically at $10°$ increments, ie $-60°, -50°, ..., +50°, +60°$), and specify known output vectors at those angles. The network is "trained" at these 13 specified angles. Furthermore, the network has, hopefully, learned sufficiently well to generalize and accurately predict output vectors for input vectors between the training points. This is called "supervised training" and it is the foundation of function approximation mentioned earlier.

Unfortunately, with multi-component input and output vectors, it is impossible for ordinary human beings to visualize function approximation in the usual sense. We can simplify the visualization of function approximation considerably if we use a scalar input variable, x, and a scalar output variable, y. At the same time, we can learn more about the radial basis function neural network. First, the radial basis functions that we use are gaussians.

$$\phi_i(x) = e^{-\frac{(x-c_i)^2}{2\sigma^2}} \tag{1}$$

for $i = 1, 2, ..., t$, for t "training" points in the input space, x. The c_i are the gaussian centers which we set equal to the x values at the training points, i.e., $c_1 = x_1$, etc., and σ is the spread of the gaussian function. The scalar output, y, is a weighted sum of the basis functions.

$$y = \sum_{i=1}^{t} w_i \phi_i(x) \tag{2}$$

Now, construct the following system of equations:

$$y_j = \sum_{i=1}^{t} w_i \phi_i(x_j) \tag{3}$$

for $j = 1, 2, ..., t$, and y_j are the known ("desired") output values at the training points x_j. We could call them sample points in our simple example. This linear system can be written as:

$$Y = \Phi W \tag{4}$$

where Y is a column vector of output values, y_j, the rows of Φ are the gaussian functions evaluated at x_j and W is a column vector having the weights, w_i, as components. If the number of training points is equal to the number of gaussians in our network (as we have assumed here), then there are as many equations as there are unknowns in the above linear system. We can then solve for the weights as follows:

$$W = \Phi^{-1}Y. \tag{5}$$

Training is complete when we solve for the weights, i.e. the network has "learned" the function. Equation 2 can then be used to find y for any value of x. This is function approximation as discussed earlier.

To further clarify the concept, we present a simple example of training using our scalar input/scalar output network. We assume that the desired output is a linear ramp function from $-60°$ to $+60°$ and the input $x \in [0,1]$. The seven training (sample) points are distributed in x as shown by the circles in Figures 1 and 2. The function is approximated with gaussian functions with centers at the training points. The approximation is shown for two values of the spread parameter, $\sigma = .01$ in Figure 1 and $\sigma = .1$ in Figure 2, demonstrating the importance of this network parameter. We use the same technique to train a network with an antenna array except the inputs and outputs are multi-component vectors.

3   The Classical Monopulse DF Algorithm

Monopulse is a classical array antenna application requiring high quality beams (low sidelobe sum beam and high slope difference beam). Of course, high quality beams require well matched elements or some sort of calibration scheme. The digital beamforming monopulse algorithm is implemented by applying two independent sets of weights to the measured amplitudes and phases of each antenna element to form two independent beams. One is a sum beam with a maximum in the direction of the source and the other is a difference beam which has a null in the source direction [5]. The sum beam is synthesized with an $\bar{n} = 4$, 30 dB Taylor aperture distribution; the difference beam uses a 30 dB Bayliss.

The algorithm takes the real part of the ratio of the difference beam to sum beam voltages and converts this value to an equivalent azimuth angle relative to the angle where the peak gain of the sum beam occurs. Because of element mismatch and noise, this predicted angle will not, in general, correspond to the true angle of the source. Sherman [5] describes these well known sources of error in monopulse angle estimation. The unavoidable error due to thermal noise in the receiver channel is given by:

$$\sigma_{\theta t} = \frac{\theta_3}{k_m} \frac{1}{\sqrt{2n(S/N)}} \tag{6}$$

for n array elements. θ_3 is the beamwidth of the sum beam and k_m is the dimensionless, normalized ratio slope and is approximately equal to 1.7 for the assumed aperture distributions for both eight and 32 element arrays. S/N is the linear signal to noise ratio at the receiver for an individual element. We use one receiver channel per element for digital beamforming[6].

The array can be partially compensated for element mismatch by taking the 32 amplitudes and phases measured at array broadside and normalizing all measured data (amplitudes and phases at each of the 32 elements at the remaining 120 angles, $-60°$ to $+60°$ in $1°$ increments excluding broadside) with these values. Exact error compensation occurs at broadside with degraded beam quality for angles off broadside. For all of our monopulse DF performance calculations, we calibrate the array in this manner.

4   Description of the Experiment

Data was taken on an X-band (10 GHz), 32 element linear array antenna to exercise our DF algorithms. The elements are separated by $\lambda/2$. The array consists of 32 columns of microstrip patch radiators, each column consisting of 8 patches combined in a phase-matched feed to provide elevation directivity. Measurements were taken in the anechoic chamber shown in Figure 3. The patches are vertically polarized and we take H-plane azimuth patterns. The antenna is phase-centered over the boom pivot point and the boom is rotated in azimuth using a stepper motor, moving the source horn around a semi-circular track from $-60°$ to $+60°$ and stopping at $1°$ intervals for measurements.

Received amplitude and phase is measured at each element in the digital beamforming tradition[6] using a coaxial rf switching matrix behind the array to connect every element, in turn, to a Scientific Atlanta 1780 phase/amplitude receiver. Thus, the horn is stopped momentarily at 121 positions, representing azimuth angles from $-60°$ to $+60°$ inclusive, and 32 phases and amplitudes are measured at each position. This data is digitized and stored in files for off-line algorithm testing.

5   Results

DF performance of the 8 and 32 element arrays is shown in Figure 4 as a function of element SNR. The solid line is Equation 6 for $n = 8$ or 32 elements. The circles represent monopulse performance using simulated data for a perfect antenna (no element mismatch and isotropic element patterns) with thermal (gaussian) noise added in sufficient quantity to give the desired SNR. Agreement with the theoretical curve is remarkable, especially since Equation 6 is derived assuming large SNR. The asterisks represent performance calculated using the monopulse algorithm on measured data for the full 32 elements. For the 8 element array, an x indicates performance calculated from measured data. The monopulse algorithm is calibrated using high SNR data as described in Section 3. Increased error using measured data is due to array element mismatch which cannot be removed with our calibration technique. The performance is still very good; the rms error is less than $\frac{\theta_3}{10}$ even for SNRs on the order of 10 dB. Since k_m is approximately the same value for both eight and 32 element arrays, and since θ_3 is inversely proportional to n, we see from Equation 6 that for a given SNR, the error scales as $\frac{1}{n^{3/2}}$.

Figure 5 shows the DF performance using the neural network. For the network, there is no theoretical curve since one has not been derived as Sherman did for monopulse. However, we use simulated data as before to obtain an estimate of this error limit; this limit is indicated by the circles

at selected SNR values. An approximate scaling law of $1/n$ for the angle error can be derived from the simulated data curves. Note that for large SNR, the performance for both arrays has a residual error of about 0.35° rms even for the simulated data which assumes a perfect array with added thermal noise only. This is an artifact of the network and is not a function of array mismatch.

6 Conclusions

The DF angle error for both the calibrated monopulse algorithm and the neural network algorithm are both quite small. For SNRs larger than about 10 dB, both produce errors less than $\frac{\theta_3}{10}$. The monopulse error is about 50% smaller than the error produced by the network for large SNR. The net's error can be reduced by further optimization of the spread parameter, σ, or a refinement of the output processing technique.

We found that for a given SNR, the rms angle error decreases as $\frac{1}{n^{3/2}}$ for monopulse and as $\frac{1}{n}$ for the network. This means that for a neural network processor, fewer elements are necessary to obtain the same angular accuracy as monopulse.

References

[1] J. Simmers and T. O'Donnell. Adaptive RBF Neural Beamforming. In *Proceedings of the 1992 IEEE Mohawk Valley Sec Command, Control, Communications, and Intelligence (C^3I) Technology & Applications Conf*, pg 94-98, June 1992.

[2] T. O'Donnell, J. Simmers, and D. Jacavanco. Neural Beamforming for Phased Array Antennas. In *Proceedings of the 1992 Antenna Applications Symposium , U of Illinois*, Griffiss AFB, NY, Sept 1992. USAF Rome Laboratory. `RL-TR-93-119, Vol I ADA268167, Vol II ADA266916`

[3] J. Simmers, H. L. Southall, and T. O'Donnell. Advances in Neural Beamforming. In *Proceedings of the 1993 Antenna Applications Symposium, U of Illinois*, Griffiss AFB, NY, Sept 1993. USAF Rome Laboratory. `RL-TR-94-20, Vol I ADA277202, Vol II ADA277203`

[4] M. J. D. Powell. Radial Basis Functions for Multivariable Interpolation: A Review. In J. C. Mason and M. G. Cox, editors, *Algorithms for Approximation*, pg 143-167, Clarendon Press, Oxford, 1987.

[5] S. M. Sherman. *Monopulse Principles and Techniques,p. 297.* Norwood, MA:Artech House, 1984.

[6] H. Steyskal. Digital Beamforming Antennas: An Introduction. *Journal of Electronic Defense*, 17(7):37-38, July 1994.
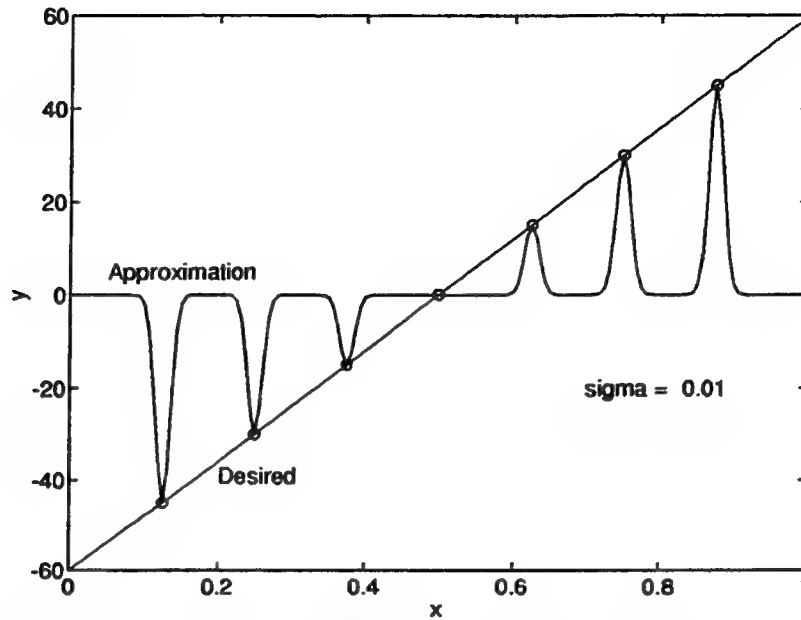
Figure 1: Demonstration of function approximation using Gaussian radial basis functions with $\sigma = .01$. The seven training points are indicated by the circles. The individual weighted Gaussians are obvious and the approximation is poor.
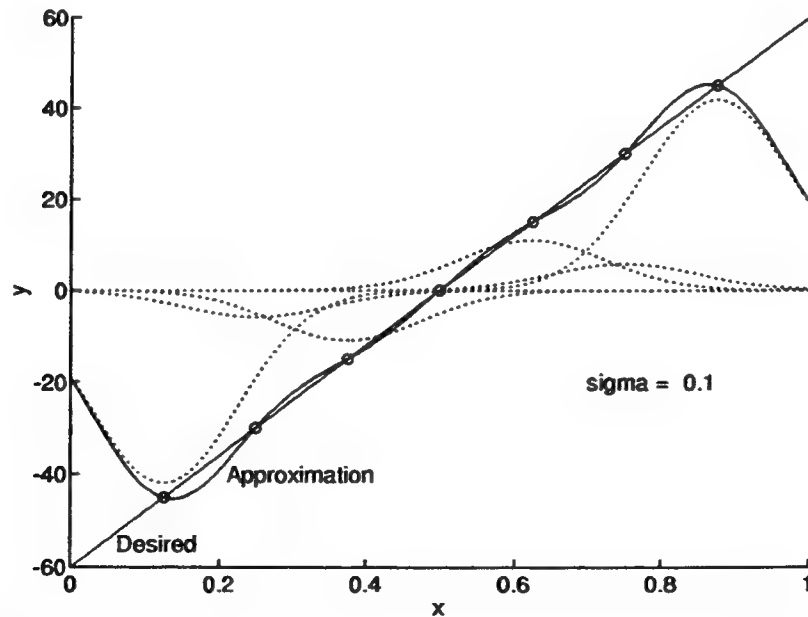
Figure 2: Demonstration of function approximation using Gaussian radial basis functions with $\sigma = 0.1$. The seven weighted Gaussians are shown dotted and the approximation, shown by the solid curve, is better.

Figure 3: A semi-circular track with a movable boom positions a 10 GHz source horn at azimuthal locations from −60° to +60° in 1° increments. Also shown is a source horn fixed in azimuth for dual-source measurements.
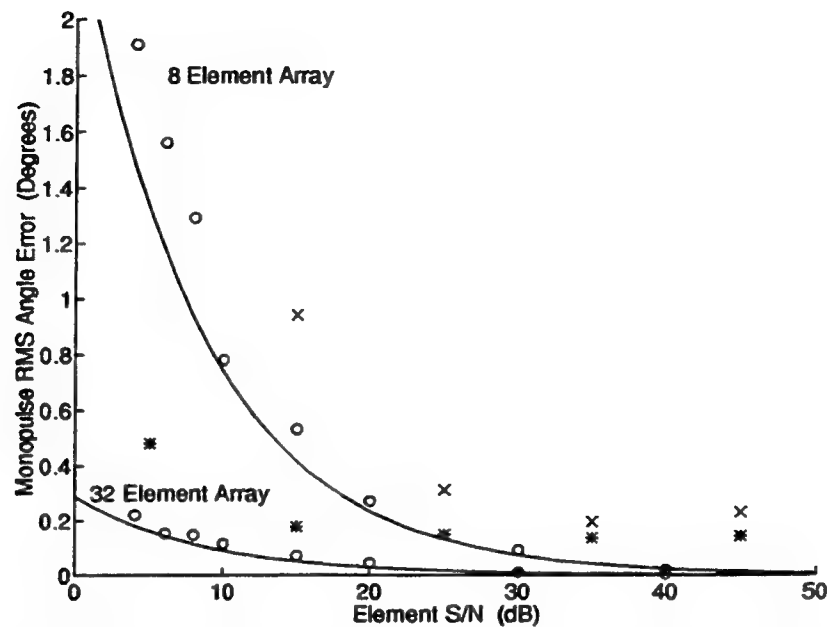
Figure 4: Monopulse DF performance for both the 8 and 32 element arrays. Circles indicate simulated data performance. Measured data performance is indicated by an x for the 8 element array and by an * for the 32 element array. The solid curves are calculated from Equation 6 for 8 or 32 elements and represent the theoretical monopulse rms error limits.
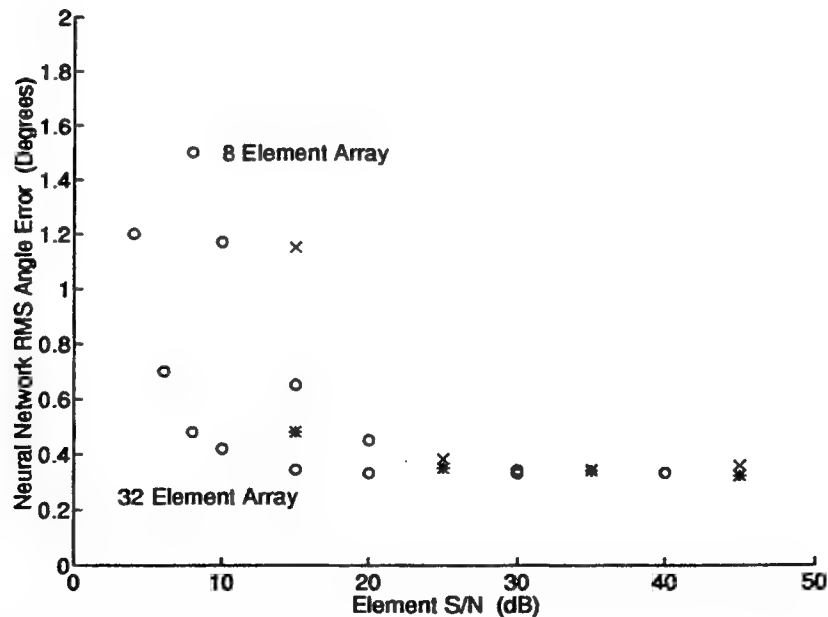
Figure 5: Neural network DF performance for both the 8 and 32 element arrays. Circles indicate simulated data performance. Measured data performance is indicated by an x for the 8 element array and by an * for the 32 element array.

A Microstrip Array Fed by a
New Type of Multilayer Feeding Network

Naftali I. Herscovici*, Nirod K. Das♦ and Josh Klugman♣

Seavey Engineering Associates, Inc., Cohasset, MA*
Weber Research Institute, Polytechnic University, Farmingdale, New York♦
Polyflon Company, New Rochelle, New York♣

Abstract: The excitation of the parallel plate mode in stripline configurations is a serious concern in the design of multilayer feeding networks for printed phased arrays. Using either probes or slots as means of power transfer between adjacent layers, this parasitic mode is always excited, and to assure its suppression a considerable effort has to be made. This paper presents a new method to solve this problem by using striplines with two different dielectric substrates on the two sides of its center strip. By using the asymmetric stripline only in the vicinity of a source of excitation, a surface-wave is excited at a significantly weaker level rather than the parallel plate mode. A sixteen element microstrip array was designed and tested to demonstrate this new concept.

1. Introduction

Aperture coupled patch arrays present significant advantages over the microstrip- or probe-fed patches:

- The crosspolarization level is very low (theoretically null in the principal planes).
- The typical effects due to the interaction between the radiating elements and the corporate feeding network(shift in the resonance frequency, undesirable radiation from the feeding network, etc.) are eliminated.

The feeding network behind the aperture-fed antenna can be further extended in many cases to a multilayer structure which include the stripline as the main

382

transmission line. The special case in which the feeding network is a single-layer microstrip circuit, was extensively treated in the literature [1,2]. There, the back radiation, even from small slots, can result in undesirable effects. Instead, a parallel-plate (either stripline or shielded microstrip) feed structure would be preferable order to avoid the back radiation, provided the excitation of the parasitic parallel plate mode can be avoided or significantly suppressed. This is sometimes done by shorting the two groundplanes in 'strategic points' on the circuits. This method, extensively used in the industry, does not fully solve the problem, and many times these posts affect the predictability of the circuit characteristics. Even when such shorting posts are used, exciting the stripline by a perpendicular probe or a slot can be impractical [3].

This paper describes the implementation of a new approach for the suppression of the parallel-plate mode in striplines. A microstrip-to-stripline-to microstrip coupler was built and tested to illustrate the idea. Finally, a sixteen element microstrip array fed by this new type of feeding network was successfully built and tested.

2. Theory

Any multilayer configuration involve striplines and sources that may excite the parallel plate mode. The parallel-plate mode is usually suppressed by electrical posts. Beyond the fabrication complexity involved, these posts create imperfect waveguides, and imperfect cavities (around slots) which might considerably affect the predictability of the circuit performances. Since slots are used instead of via holes as a mean of power transfer between adjacent layers, the use of shorting posts makes the use of the slots pointless. One way to avoid the posts would be by using asymmetric striplines (the two substrates are of different dielectric constants). This solution is acceptable as long as coupling to both sides of the stripline is not necessary. However, in multilayer configurations the coupling to both sides is needed, and the side having the lower dielectric constant will have a lower coupling level.

383

The new solution proposed here involves the use of an asymmetric stripline only in the vicinity of the coupling slot. In this case, since the main source for excitation of the TEM mode is the slot, a small area surrounding the slot (a small plug) is designed with a higher dielectric constant than the rest of the circuit. Thus the area surrounding the slot would appear to be closer to a microstrip line than a stripline, considerably reducing the excitation of the parallel mode (Figure 1). The discontinuity between the area surrounding the slot and the rest of the circuit, however, excites the TEM parallel plate mode, but at a much lower level.
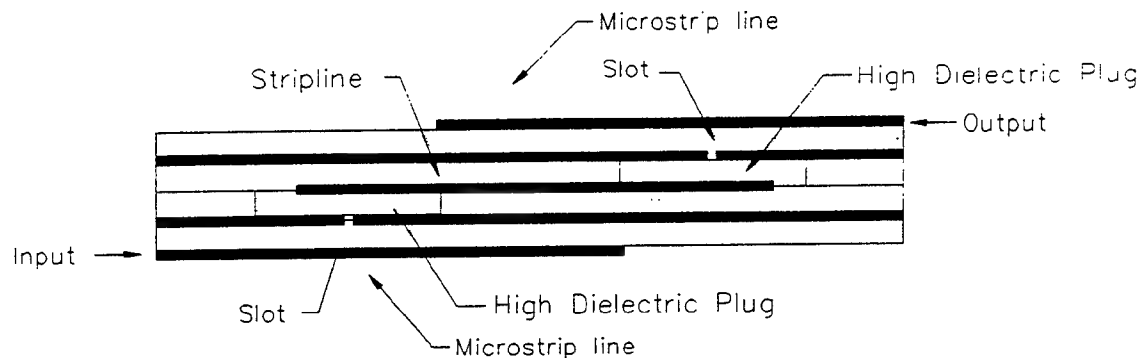
Figure 1 - The excitation area of a slot fed asymmetric stripline
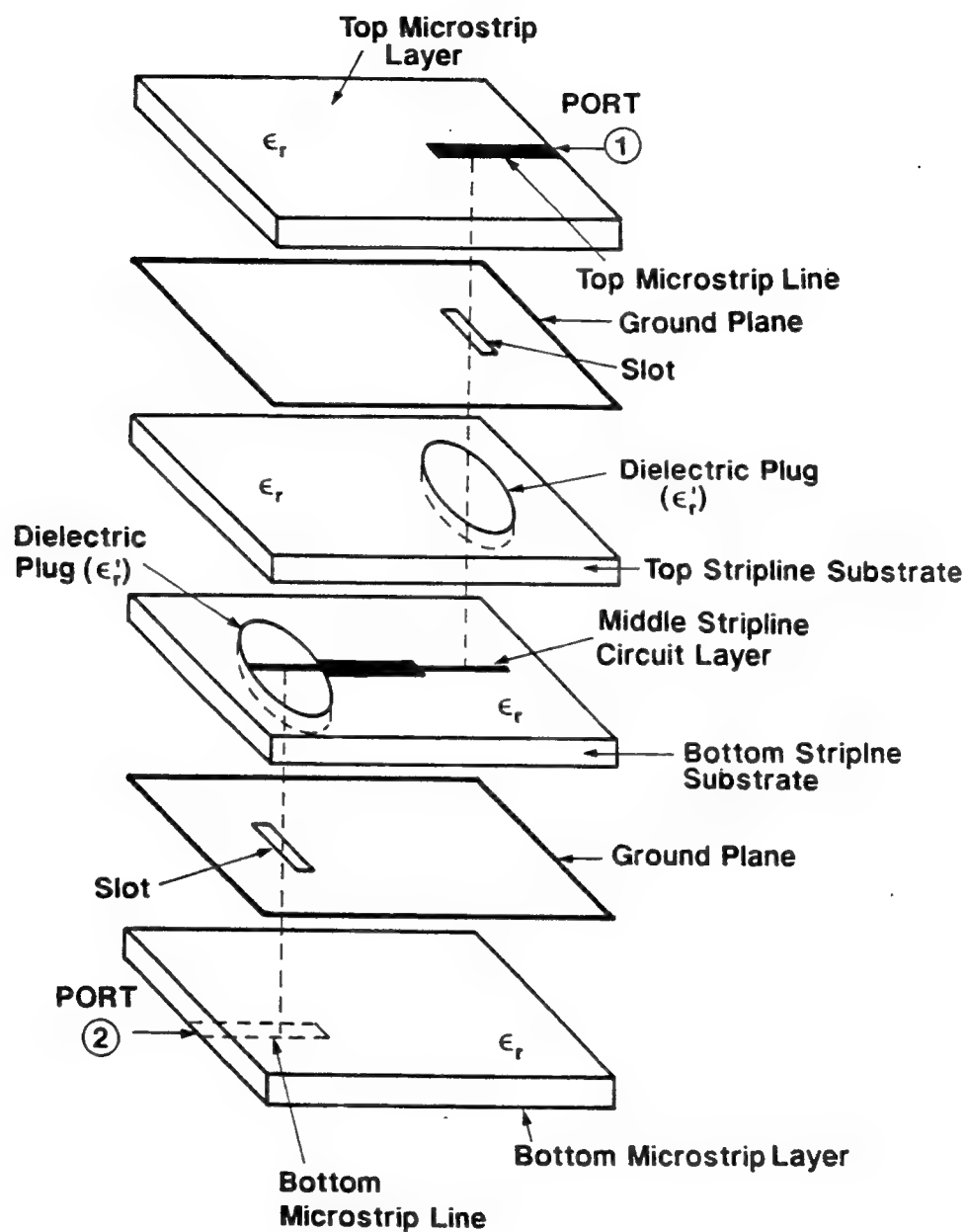
Figure 2 - The non-homogeneous multilayer coupler.

3. A non-homogeneous asymmetric coupler

The non-homogeneous two-port coupler shown in Figure 2 was built and tested. This multilayer configuration consists of a microstrip line, coupling to the stripline through a slot in the common groundplane. The other end of the stripline couples to a second microstrip line through another slot printed on its top groundplane. This is a two-port circuit, with all other transmission line ends terminated in stubs.
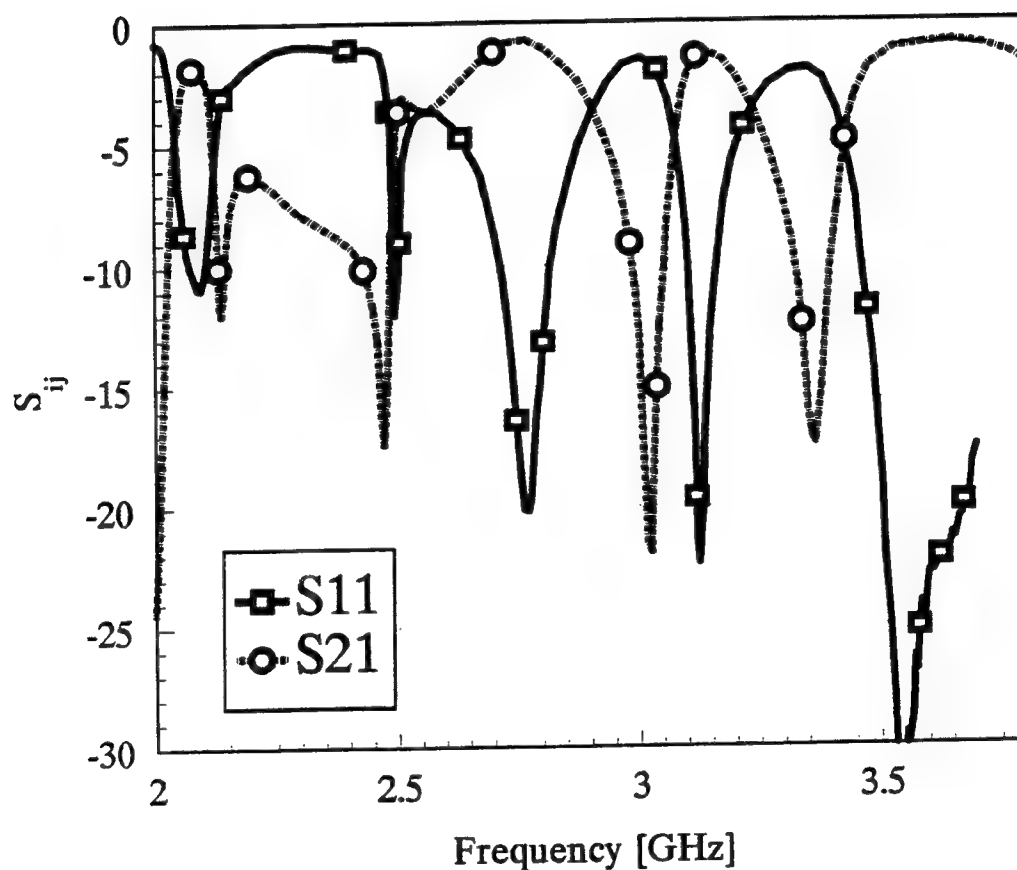
Figure 3 - The measured insertion loss and the return loss of the coupler shown in Figure 2.
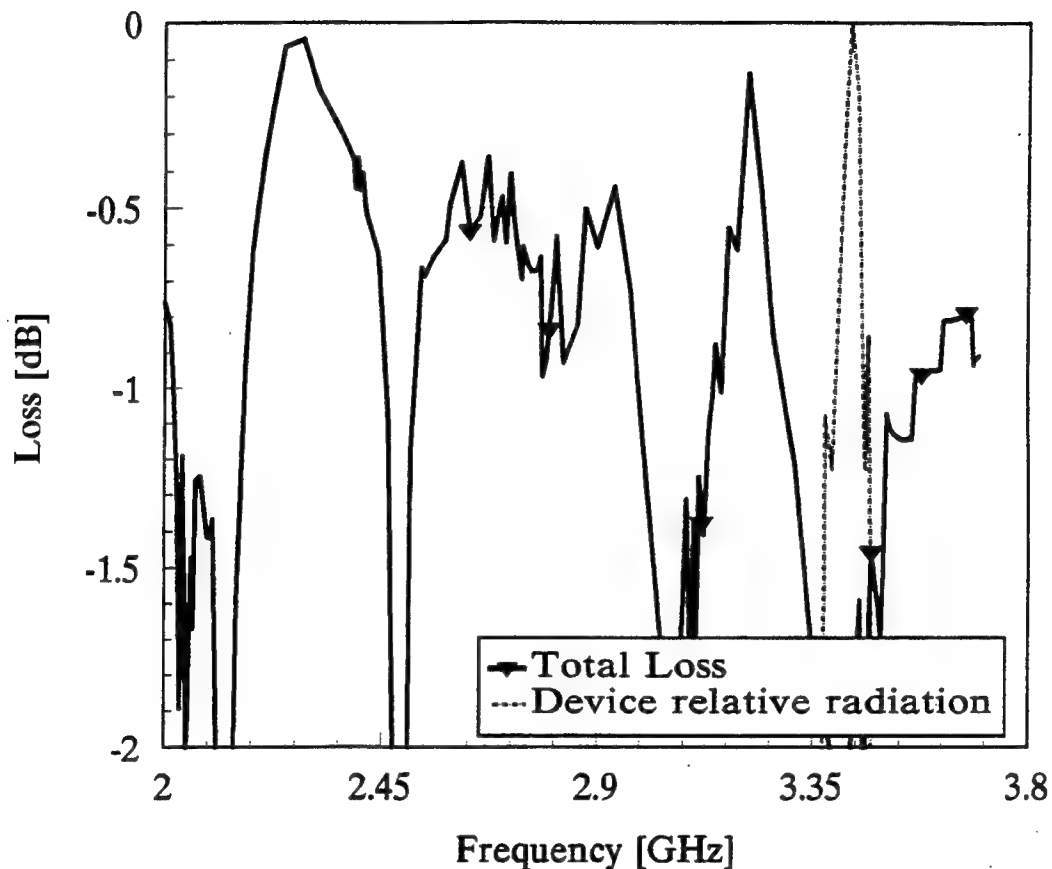
Figure 4 - The total losses of the coupler.

All lines are designed with a 50 Ω characteristic impedance, and the substrates are 31 mils Polyflon Cuflon (ε_r =2.1). The coupling slots are 2.4 cm by 0.1 cm. The microstrip as well as the stripline stubs are 1.5 cm long. The plug insert is 3 cm in diameter and 25 mils in height. The remaining gap between the plug and the groundplane (6 mils) was filled with bonding film (ε_r =2.5). Figure 3 shows the

measured insertion loss and return loss of the coupler. Figure 4 shows the total losses of the coupler computed as,

$$Loss = |S11|^2 + |S21|^2 \tag{1}$$

These losses include dielectric losses, conductive losses, surface wave losses (for microstrip) the slots radiation and power lost to the parasitic parallel plate mode (for stripline). The first three type of losses can be easily computed and subtracted from the total loses ($\cong 0.3$ dB). These losses are inherent and depend only on the type of the material used and should be excluded in the evaluation of the coupler performances. The ability to control the other losses makes this configuration different from other multilayer geometries. The impact of the slot radiation can be learned simply from the measurement. Figure 5 shows the broadside radiation of one side of the coupler. As it will be shown, the slot is not the only source of radiation. Across the 2 to 3.8 GHz, Figure 4 shows a number places where there is a significant loss. The comparison of the two graphs in Figure 4 and 5 allows us to discriminate between the sources of these losses (Table I):

Frequency Band	Main source of loss	Notes
2.5 GHz	Parallel plate mode.	The S_{11} shows a good match. There is no evidence of strong radiation.
3.1 GHz	Device radiation.	See Figure 5.
3.4 GHz	Slot radiation.	See Figure 5.

Table I
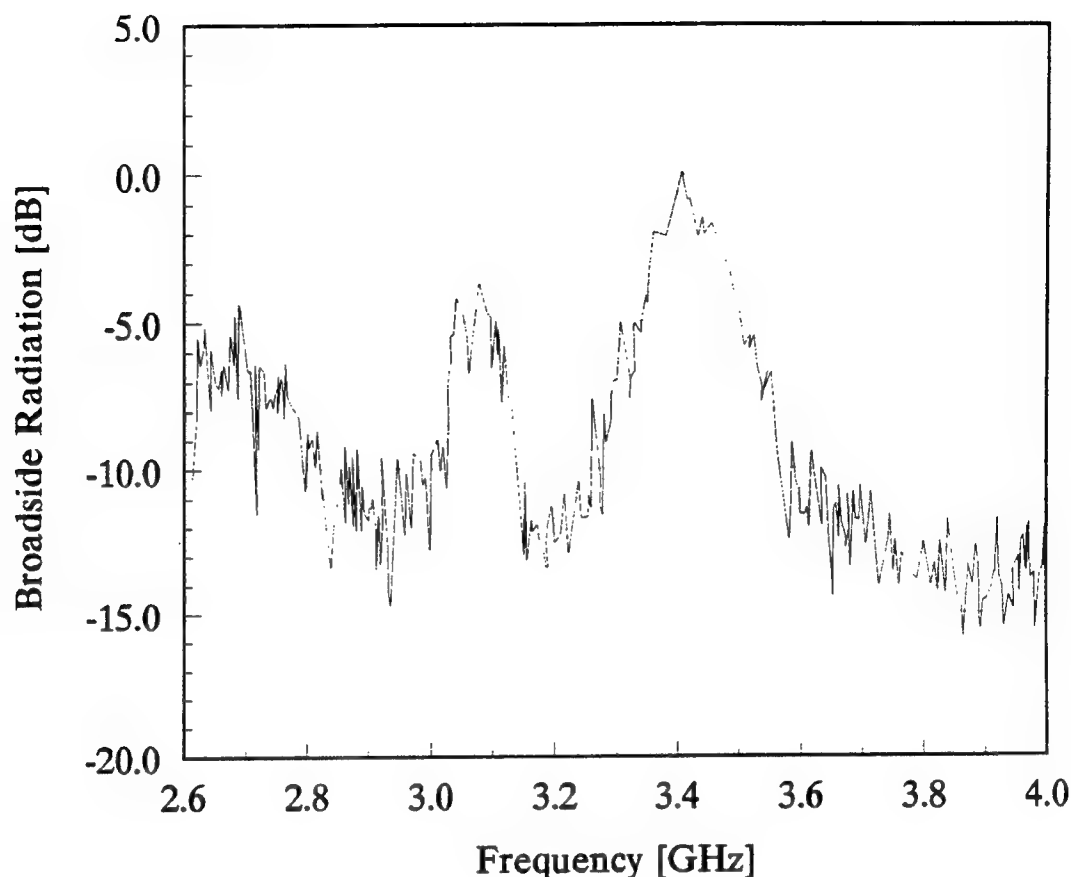
Figure 5 - The broadside radiation of one side of the coupler

In addition, some radiation is noted between 2.6 GHz and 2.8 GHz. This radiation is attributed to the stripline which is 6 cm long. Around these frequencies, it acts as a printed dipole.

Overall, two frequency bandwidths (which show a very low level of losses) are of interest: 2.18- 2.45 GHz and 3.5- 3.8 GHz. The first one exhibits very low losses and the matching of the two ports is expected to produce an insertion loss, better than 0.5 dB. In the second band, the coupler is very well matched already (S_{11}<-15 dB) and considering the dielectric and ohmic losses, the insertion loss is at an acceptable level.

4. The 4×4 array

Finally a 4×4 elements microstrip array was built. All patches are 2.3 cm by 2.3 cm and are printed on a 31 mil Polyflon Cuflon (ε_r=2.1). Each patch is fed by a slot in the common ground-plane 1 cm long and 0.1 cm wide. The feeding network is stripline made of two 31 mil layers of Polyflon Cuflon. The slot areas are replaced by high dielectric plugs (Rogers 6010, ε_r=2.1). The diameter of the plug is 2 cm. Since the commercial available 6010 is 25 mils thick (or thicker than 31 mils) the 6 mils gap was filled with Teflon bonding film. The stripline feeding network is fed through a slot by 50 Ω microstrip line printed on a 31 mil Polyflon Cuflon. The stacked layouts of the antenna are shown in Figure 6.
The radiation patterns of this array are shown in Figure 7. A crosspolarization level of more than 30 dB is noticed in the H-plane and about 26 dB in E-piane. Another 4×4 microstrip array fed by a corporate microstrip feeding network was also built to evaluate the efficiency of the multilayer array. At resonance, the gain of the multilayer array is smaller by about 1 dB.

5. Conclusions

Using the high-dielectric plugs instead of via-holes eliminates the need for ohmic contact which sometimes is difficult to insure in the vias. This technology also requires one less plating round which considerably reduces the price of the product. All the layers are simultaneously processed and bonded together. A 4×4 microstrip array fed by a feeding network which makes use of this new technology was built and tested.

REFERENCES

[1] D.M.Pozar, "A Reciprocity Method of Analysis for Printed Slot and Slot-Coupled Microstrip Antennas", IEEE Trans. on Antennas and Prop., Vol. AP-34, No. 12, pp.1439-1446, December 1986.

[2] P.L.Sullivan and D.H.Schaubert, "Analysis of an aperture coupled microstrip antenna", IEEE Trans. on Antennas and Prop., Vol. AP-34, No. 8, pp. 977-984, August 1986.

[3] N.Herscovici, "Analysis of Aperture Coupled Microstrip Transmission Lines and Antennas ", Doctor of Philosophy Thesis, Department of Electrical and Computer Engineering, University of Massachusetts at Amherst, September 1992.
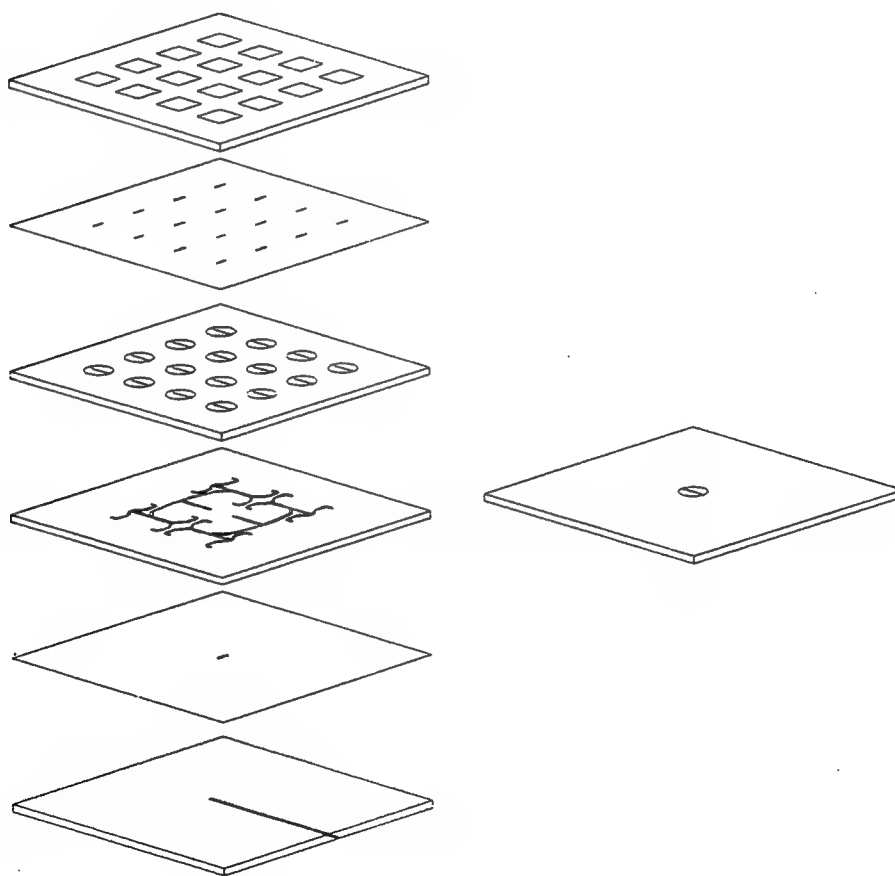
Figure 6 - The stacked layout of the multilayer antenna.
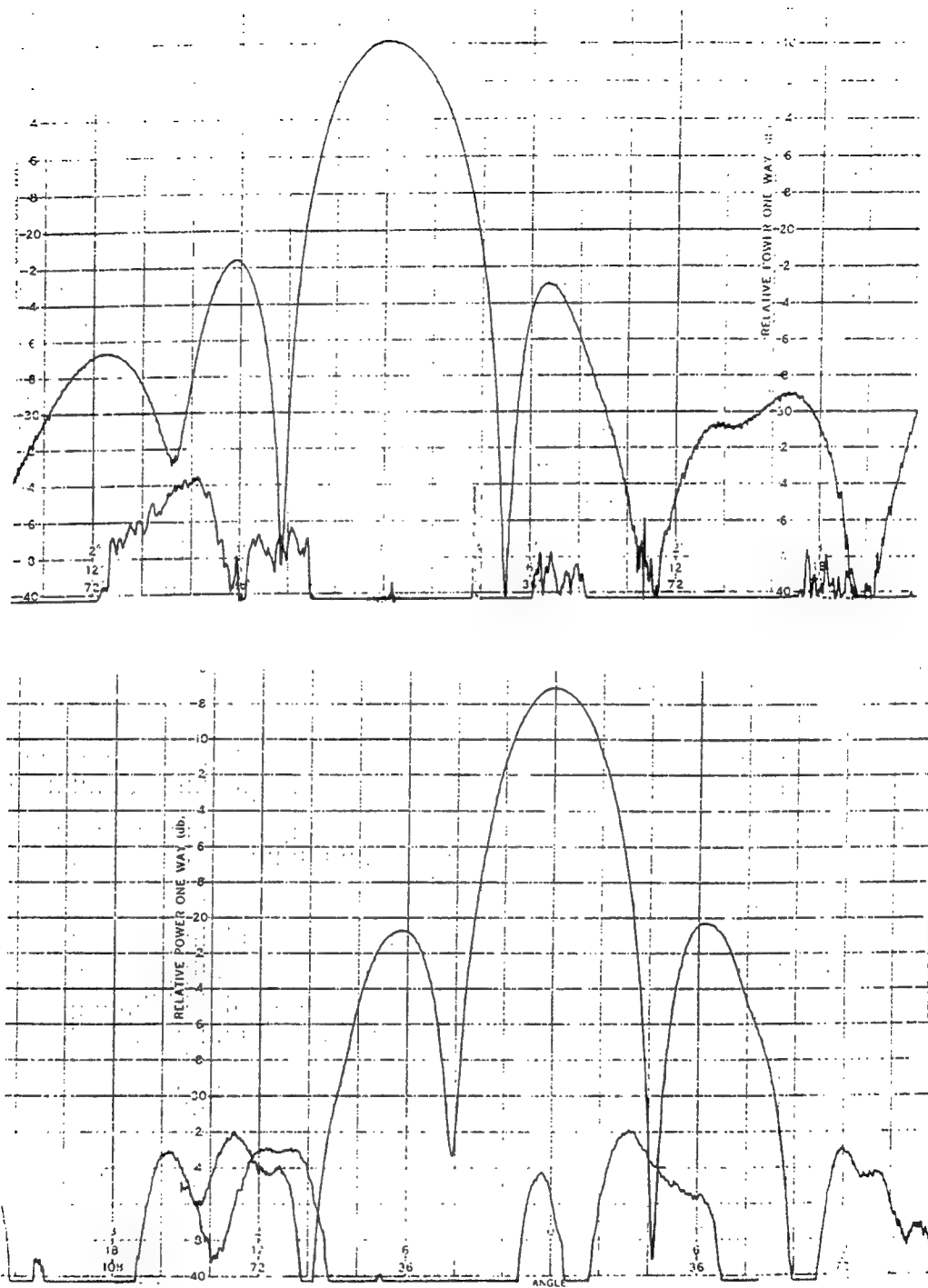
Figure 7 - The radiation patterns of the antenna shown in Figure 4.
a. H-plane
b. E-plane

Cust. P/N

REF 2

Seavey P/N

2.

S/N _____

Date _____

Technician

_____

Inspector

_____

Spec. Ref.

_____

STRT +4.2000GHz   CRSR +4.4549GHz   STOP +4.8000GHz

SEAVEY ENGINEERING ASSOCIATES, INC.
ANTENNA DESIGN AND DEVELOPMENT

HEWLETT-PACKARD SCALAR NETWORK ANALYZER SYSTEM 8756A/S

Return loss_____at ports_____
Insertion loss_____between ports_____and_____

Report No. ____
Figure No. ____
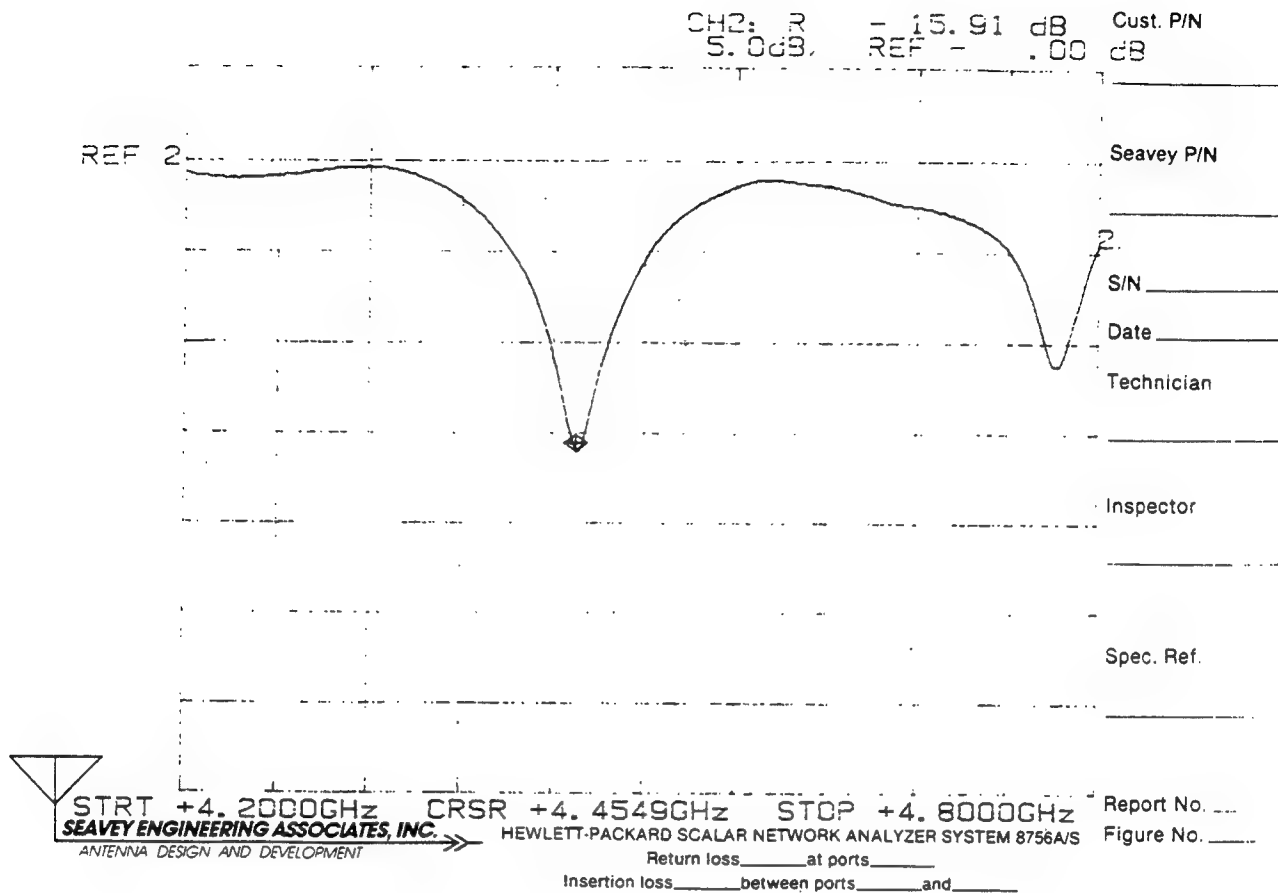
Figure 8 - The return loss of the antenna shown in Figure 4.

MULTI-BEAM FEED FOR PHASED ARRAY ANTENNA

Harold Shnitkin

Westinghouse Norden Systems, Inc.
P.O. Box 5300
Norwalk, CT 06856

ABSTRACT

This paper discusses the design of a single antenna aperture capable of very rapid switching between different antenna pattern shapes and beam directions. If focuses on the design of a microwave network or feed configuration for a 1-D, E-scanned array which selectively switches one of several transmit/receive ports to all the radiators of a phased array under the control of the electronic phase shifters used for beam scanning. Each port corresponds to a different antenna pattern, such as monopulse/sum/difference, new beam pointing direction in space, cosecant shaped beam, or multi-port interferometer with very low sidelobes.

Methods for designing monopulse and multi-direction beam apertures are well known (as per Lopez, Blass, and Butler). This paper mainly addresses the design of a microwave network, similar to a Blass network, capable of electronic witching between various pattern shapes and aperture illuminations.

The theory of the phase shifter controlled microwave network, its methods of implementation, and practical techniques for electronic trimming of the aperture illumination to compensate for manufacturing tolerance errors will be presented. In addition, a computer program to design the network and to determine component parameters, losses, isolation between ports (e.g., modes), and resulting accuracy will be explained.

Finally, the application of the above concept to the Joint STARS antenna, which switches electronically between a low-sidelobe, full aperture pattern and three, low-sidelobe aperture, interferometer patterns, will be shown.

1.0 Background

Multi-beam linear array feeds which contain multiple antenna ports to allow a single aperture to generate several different radiating beams have been around for several decades and were invented by Blass, Butler, and Lopez. (REF. 1, 2, 3) These

feeds demand orthogonality between beams as an essential criterion in order to avoid serious interference and perturbation between the different beams and/or their corresponding feed networks. This interference originates from the coupling of one beam's feed to the next and the subsequent re-radiating of this coupled energy; the latter causes antenna pattern errors and, furthermore, channels energy into the other beam's feeds and their loads, resulting in high insertion loss.

Orthogonality between feeds (or beams) is achieved mainly by beam separation in space, in such a manner that every beam maximum spatially coincides with the nulls of the radiation patterns of all the other beams. This means, for example, a minimum of $1.2 \times$ beamwidth separation and a -3.9 dB cross-over between adjacent beams. Needless to say this criterion is automatically observed in the case of one sum and one difference pattern. However, in other applications orthogonality imposes a beam-pointing/beam-separation limitation on any set of multiple beams, while it permits independent and simultaneous operation of all beams with one single aperture.

2.0 Multibeam Feed Network for E-Scanned Antenna

When a multi-beam feed is used in conjunction with a one-dimensionally E-scanned antenna, the following additional advantages are realized, provided the simultaneous use of all the different beams is sacrificed :

1. Beams can be pointed in any direction
2. Beams can be individually selected via the scanning phase shifters

3. Phase errors introduced by the inter-feed coupling, by manufacturing toler-
ances, and by the adherence to the orthogonality criterion can be completely
compensated with the scanning phase shifters.

Thus the designer remains at liberty to select any beam direction and any beam shape! This allows, for example, switchable multi-beam aperture/feed designs with (1) several co-directed, but differently shaped beams, (2) a series of very closely spaced beams, and (3) several co-directed, low sidelobe beams of different beam widths for multi-port interferometer application. In all cases the scanning phase shifters are used to re-position the beams to overcome the beam separation require-ment within the multi-beam feed as dictated by the orthogonality criterion, mentioned above.

The multi-mode feed network is essentially an RF switch, controlled by a row of aperture phase shifters; only one antenna input port at a time is effectively connected to all the radiators and provides its unique aperture excitation for the desired beam shape while, at the same time, only a negligible amount of energy is coupled into the unused feeds.

3.0 Circuit Analysis

To analyze the circuit of a multi-beam feed, a particular two-mode feed configu-ration has been selected, namely, a full aperture, E-scanned, low sidelobe mode plus a three one-third aperture, space-coincident, E-scanned, low sidelobe "interferom-eter" mode. This configuration is borrowed from the Joint STARS radar antenna, illustrated in Figure 1 and schematically shown in Figure 2. (Details in REF. 4) For full-aperture mode operation, the 12-way power divider shown distributes RF to
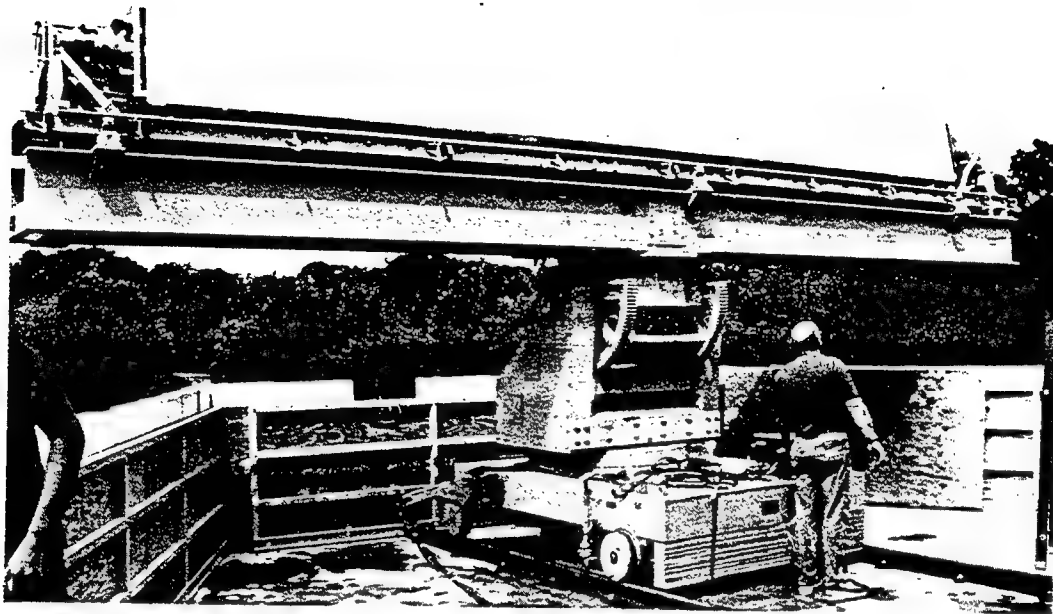
Figure 1. Joint STARS Antenna Aperture

twelve 38-way dual mode feeds (DMF ÷ 38's), from whence 456 radiators are excited via phase shifters and compensating delay cables. For interferometer mode operation, each of the three 4-way power dividers shown distribute RF to four DMF ÷ 38's, from whence 456 radiators are again excited.

By appropriate commands to the 456 phase shifters, the twelve dual-mode feeds convert one total, low-sidelobe aperture with tapered illumination for the mapping mode into three low-sidelobe, one-third apertures for the interferometer/tracking mode, as shown in Figure 3. (Details in REF. 4) Each 38-way dual mode feed (DMF ÷ 38) consists of two cascaded, directional coupler, travelling wave distribution feeds separated by 38 delay sections, as shown in Figure 4. The dual mode feed resembles a two-beam Blass-type feed; but, instead of the usually similar aperture illuminations and different linear phase tapers of the various beams of the Blass feed,
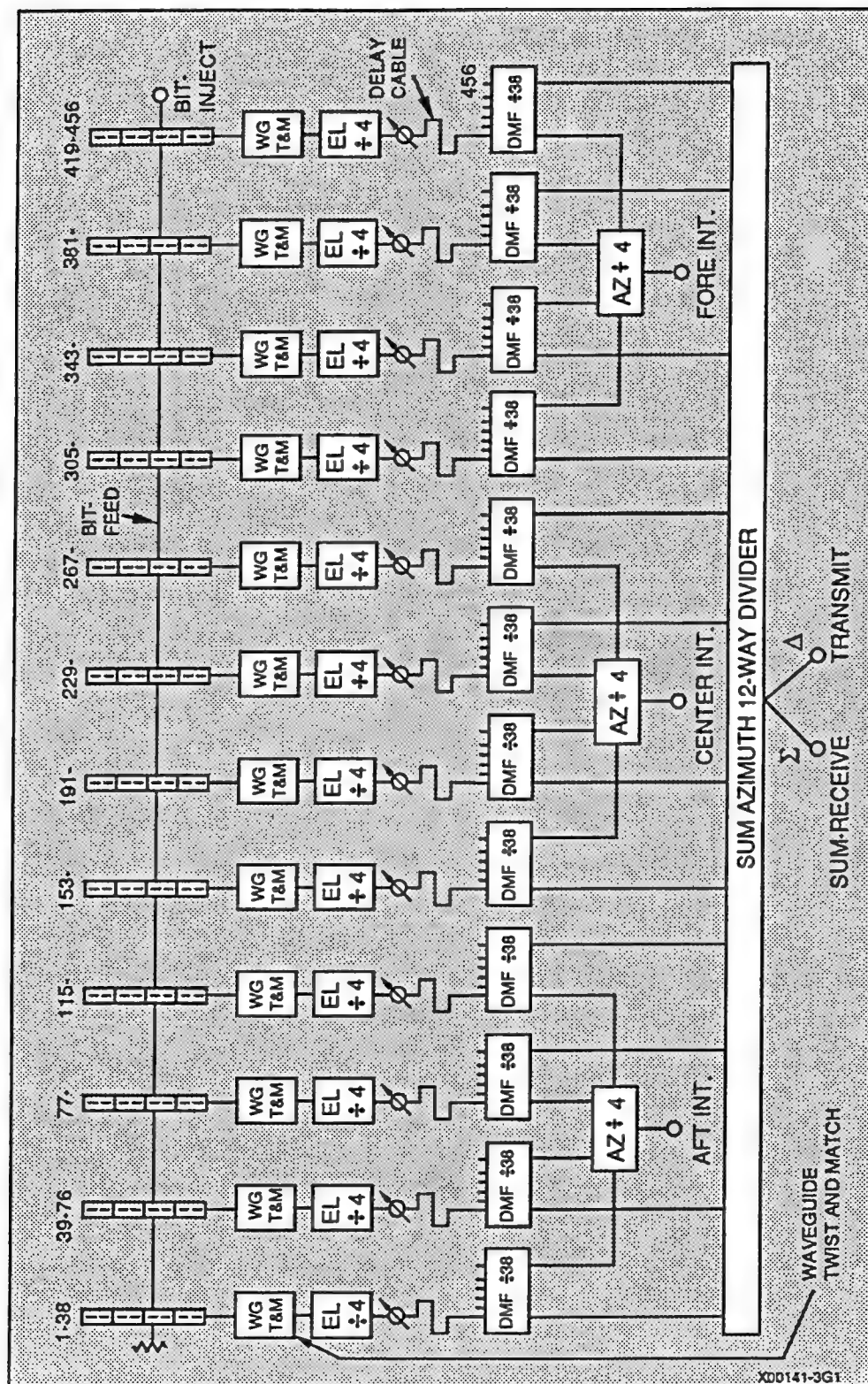
Figure 2. Antenna System Block Diagram

SUM - FEED AMPLITUDE TAPER

FORE	CENTER	AFT

INTERFEROMETER - FEED AMPLITUDE TAPER
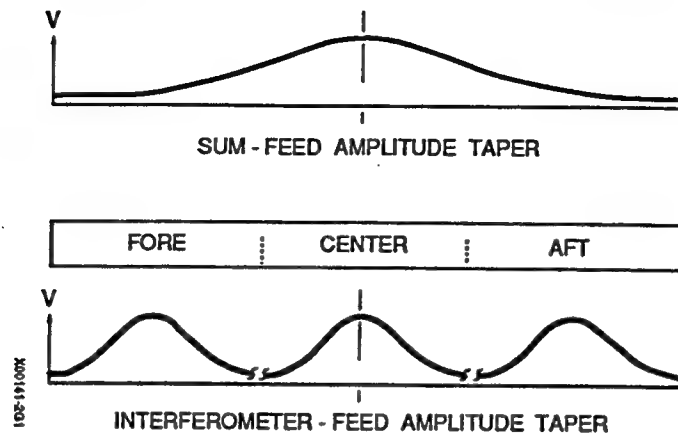
Figure 3. Aperture Illumination Function

INTER-FEROMETER

WAVEGUIDE DIRECTIONAL COUPLER

U.S.P. 4,348,679
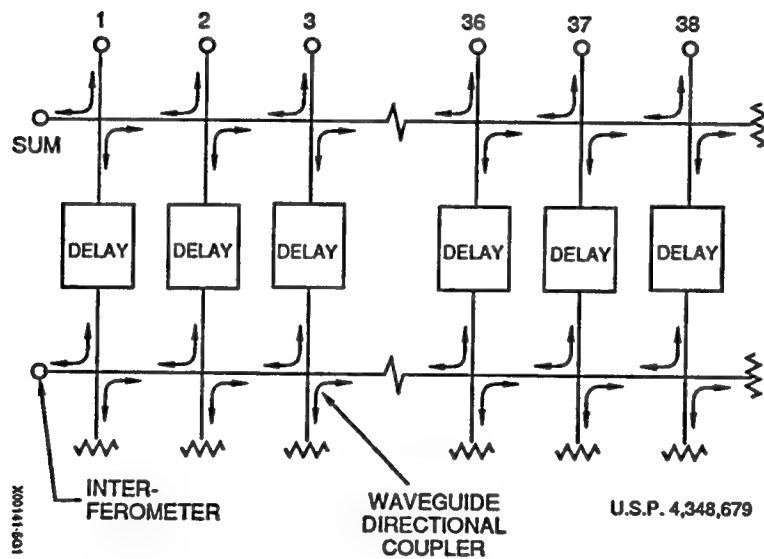
Figure 4. Dual Mode Feed (DMF ÷ 38) Schematic

two distinct illuminations and two phase functions, specifically designed for maximum decoupling between the two travelling wave feeds, are generated. Mode switching by means of the phase shifters is accomplished by selectively adding one of two available phase collimating functions to the scan angle phase taper value. Each

phase collimating function produces a linear aperture phase taper for only *one* mode and a de-collimated, and thus decoupled, aperture phase for the *other* mode. These phase collimating functions are obtained from an aperture alignment technique, described in REF. 4.

The RF delays between the two rows of couplers provide the required orthogonality (i.e. isolation) between the two modes (or feeds) as well as a vernier adjustment of the amplitude provided to each radiator. Their operation will be described in detail in the next section.

4.0 Detailed Analysis of the Feed

Since the upper or sum feed in Figure 4 is directly connected to the aperture and therefore not perturbed by any other feed, it can therefore be designed in the conventional manner. The lower or interferometer feed, however, is perturbed by the presence of the upper feed. When the interferometer port is energized, the energy flowing upward in the 38 vertical lines is intercepted by the upper feed, causing energy to flow to the right toward the upper terminating load. At each coupler some of that energy will be re-radiated by the aperture and the remainder terminated in the load at the right. This is illustrated by the RF paths shown in Figure 5. The re-radiation explains the perturbation of the aperture excitation while the upper load dissipation gives rise to additional insertion loss. To minimize these two detrimental effects, the orthogonality principal need to be observed.

To best understand the orthogonality principal, one should examine the dual mode feed in the receive mode and think of each single beam feed as a "space filter"

400
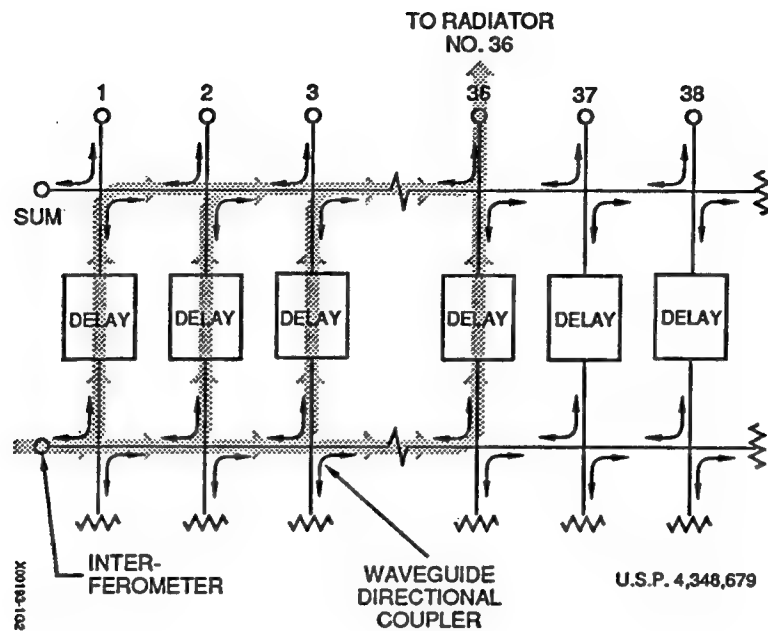
Figure 5. Signal Flow in Dual Mode Feed

tuned to a specific scan angle in space. If the received RF energy matches the phase taper of *that* scan angle, it will exit at that beam's port with the efficiency of that feed (usually > 90%). As the received phase taper is altered from the above value, the exiting energy will vary in amplitude just as the beam's antenna radiation pattern. At specific phase tapers, a pattern null is encountered and no energy will exit; thus the energy intercepted by that feed will be zero. Aiming the lower feed beam into the null direction of the upper feed beam therefore guarantees complete beam isolation or *orthogonality*. However, some residual coupling between the various vertical radiator lines of Figure 4 will still occur and create a minor aperture perturbation. This then is compensated by a change in each coupler's coupling value in the lower feed.

401

5.0 Practical Considerations

The feed design starts with a range of practical coupling values (between 7 and 25 dB) for both, the upper and the lower feeds, computed as a conventional, unperturbed, travelling wave feed. Subsequently, to assure adequate system bandwidth, a null of the upper feed's radiation pattern is selected as far away from the peak of the beam as possible, so that the pattern null is flanked by very low sidelobes. The corresponding inter-feed phase delays are then computed from the phase taper of *that* null position and the intra-feed phase delays between the adjacent radiators.

In order to optimize the coupling values of the lower row of directional couplers of the dual mode feed of Figure 4, a computer program, which calculates the 38 signals exiting to the radiators has been written. This program assumes a known value of input power, a given set of coupling values of the couplers of both feeds, and a set of 38 inter-feed phase delays. It also takes into account all coupling paths, waveguide losses, and coupler insertion phase, but assumes infinite coupler directivity and unity VSWR. Finally, it performs a comparison between the computed and the required aperture excitation. This computer program is used to calculate the resulting aperture excitation. Should the deviation between the computed and required excitation values exceed the acceptable tolerance, the computer can be asked to reset any coupling value to remove that deviation. Alternatively, the operator may select another value of inter-feed delay phase or coupling value to correct for his deviation. Both types of corrections of course must be done for one coupler at a time, in sequence, starting at radiator No. 1 (left).

Finally, because of errors arising from VSWR, from the couplers' finite directivity, and from manufacturing tolerances, the assembled dual mode feed will exhibit small residual errors in aperture amplitude values. These are removable by means of a vernier adjustment of the inter-feed delay sections. (Lockable screwdriver adjustments!) This adjustment consists of a change in the phase of the signal supplied by the lower row coupler relative to that by the corresponding upper row coupler; (see Figure 5) since the total signal delivered to each radiator is their vector sum, the inter-feed phase adjustment translates into an output amplitude vernier adjustment. Provided that the signal in the upper feed is at least one-tenth the signal in the lower feed and that the relative phase of these two signals is neither zero nor 180 degrees, amplitude adjustment by means of delay phase change is always possible. The computer program also permits monitoring of the upper feed signal level to assure that the phase and amplitude conditions, mentioned above, are always met and to ascertain the amount of energy absorbed in the load. In practice, the vernier adjustment via phase trimming must proceed in sequence, again starting from coupler No. 1.

6.0 Implementation

The Joint STARS application, as show Figures 1 and 2, contains twelve, 38 radiator, dual mode feeds. Figure 6 shows a single dual mode feed assembly, manufactured by *Atlantic Microwave* and Figure 7 is a photograph of several assemblies installed in the Joint STARS antenna configuration.
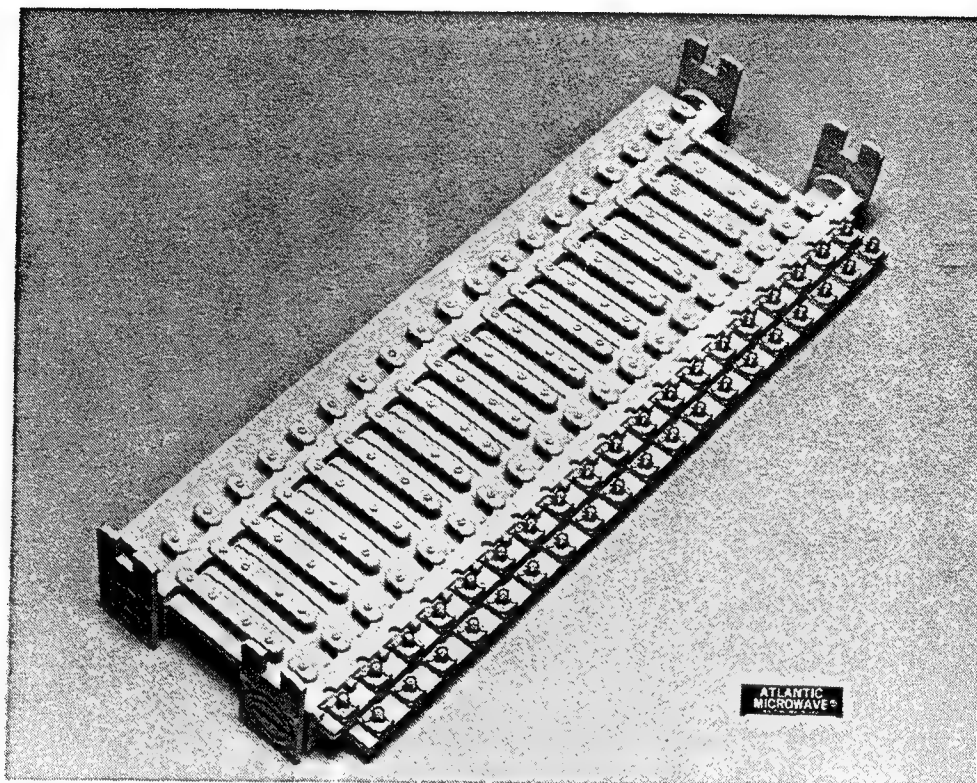
Figure 6. Production Configuration Dual Mode Feed

7.0 References

[1] J. Blass, "Multidirectional Antenna, A New Approach to Stacked Beams", *IRE National Convention Record*, Part 1, 1960, pp. 48-50

[2] J. Butler, "Multiple Beam Antenna", Sanders Assoc., Nashua, NH, Memo: RF-3849, January 1960

[3] A. R. Lopez, "Monopulse Network for Series Feeding an Array Antenna", *IEEE Transaction*, vol. AP-16, No. 4, July 1968

[4] H. Shnitkin, "A Unique Joint-STARS Phased-Array Antenna", *Microwave Journal*, January 1991, pp. 131

REAR-VIEW OF ANTENNA

DUAL MODE FEEDS

DELAY CABLES

PHASE SHIFTERS

4-WAY DIVIDERS
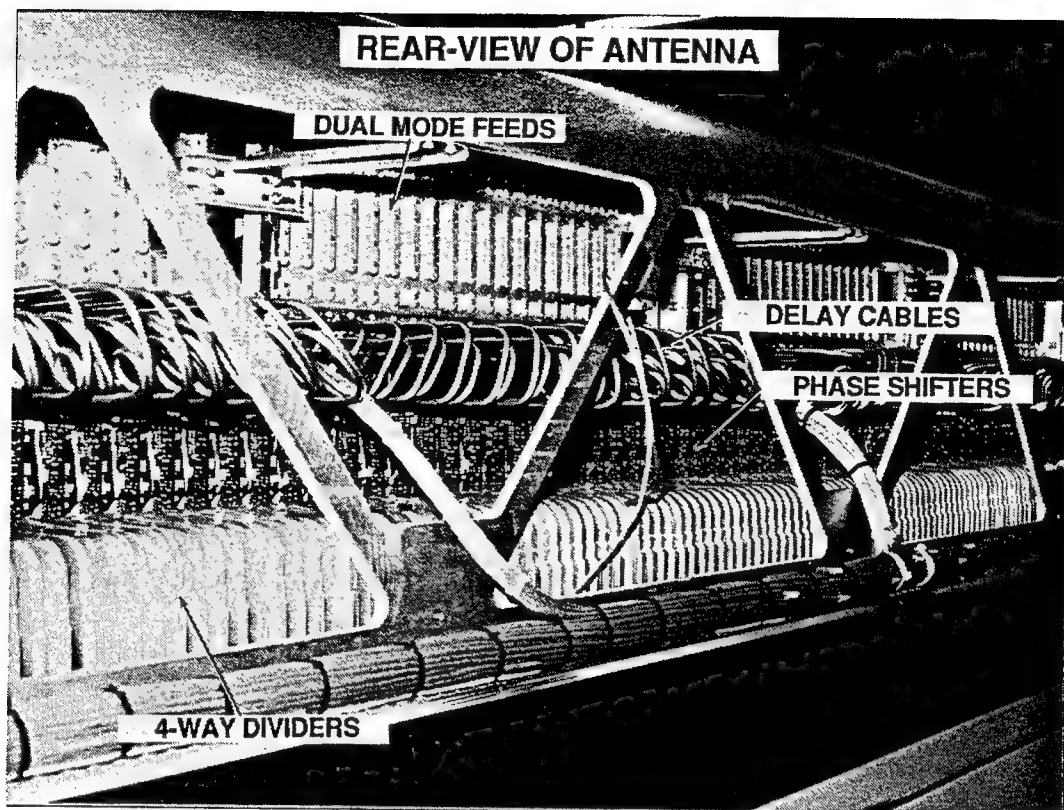
Figure 7. Joint STARS Antenna

AN ADAPTIVE MATCHING CIRCUIT FOR PHASED -ARRAY RADARS

Christopher Lyons, Wes Grammer, Peter Katzin, Mitchell Shifrin

HITTITE MICROWAVE CORPORATION

ABSTRACT

An S-Band high power adaptive matching network capable of sensing and correcting for arbitrary phase load mismatches up to 5:1 has been demonstrated. Settled VSWR's of less than 2:1 in 500 nS is typical. The circuit was designed to handle up to 10 Watts CW by utilizing novel series stacked MESFET switch MMIC's capacitively coupled to a microstrip transmission line.

1.0   INTRODUCTION

Large variations in antenna load impedance, caused by beam steering [1] can adversely effect output power, gain and efficiency characteristics in T/R modules. Traditionally, a ferrite isolator is inserted between the power amplifier and the radiating element to isolate the amplifier from the high VSWR's and to provide a constant matched load. One major disadvantage associated with the use of an isolator is that under high antenna mismatch conditions, the reflected power is absorbed by the terminated port of the isolator, thus reducing T/R module efficiency.

Increased demand for improved power added efficiency in T/R modules has generated interest in developing an efficient replacement for the isolator. One such

alternative is a tunable matching circuit. A low loss reactive matching circuit that eliminates the high antenna mismatch rather than dissipating the reflected power would result in increased T/R module efficiency. Matching circuits with external electronic control have recently been demonstrated using PIN diodes [2]. These matching circuits, however, require rigorous calibration and an increase in the number of data bits sent to each module. In addition, Pin diode switches operating at 10 Watts CW require high voltage and substantial current from the driver circuit. Such large voltage swings result in significant design complexity, power consumption, bulk, and longer response times.

This paper presents a new method for reducing the high VSWR's in T/R modules caused by beam steering without the increased T/R module interface complexity problems previously demonstrated with manually controlled tunable matching circuits. This new technique is based on an adaptive matching network that electronically senses the VSWR, and adaptively compensates for the mismatch.

The adaptive matching network hardware was implemented using a unique power conserving approach. Series stacked FET's are used in lieu of the PIN diodes. The FET switches drain negligible control current, and by stacking several switches in series, it is possible to handle power well in excess of 10 Watts, [3,4] using control voltage swings less than 10 volts.

In this paper we first present the overall system block diagram of the adaptive

407

matching network followed by a detailed description of the tunable matching circuit. Next, the power handling capability of the FET switches and physical implementation of the matching network is presented. Finally, we present photographs and test results.

2.0 BLOCK DIAGRAM DESCRIPTION

A block diagram of the adaptive matching network is shown in Figure 1. The tunable matching circuit is realized using FET MMIC switches capacitively coupled to a high impedance transmission line. With all switches off, the matching circuit appears as a 50Ω artificial transmission line running from the antenna port to the transmit port. The switches are individually selected to tune the matching circuit, over the range of possible mismatches present at the antenna port. The voltage standing wave pattern on the transmission line is monitored using four equally spaced loosely coupled detector diodes D1 through D4. The differential amplifiers and comparators convert these detected voltages into a tuning word for the digital state machine which attempts to calculate the exact switch location that will compensate for the mismatch on the antenna port. This is accomplished using a behavioral search algorithm implemented in a programmable array logic chip. The decoder / driver provides the appropriate Gate control voltage to turn the FET's on or off. A dual

directional coupler located between the transmit port and the matching circuit is used as a VSWR detector. It measures the difference between the incident and reflected powers using detector diodes (DT and DR), and issues a command to the state machine to continue searching for a more optimum matching location if the VSWR exceeds a preset threshold. Each circuit decision, and switch selection attempt is synchronous with an external clock oscillator.
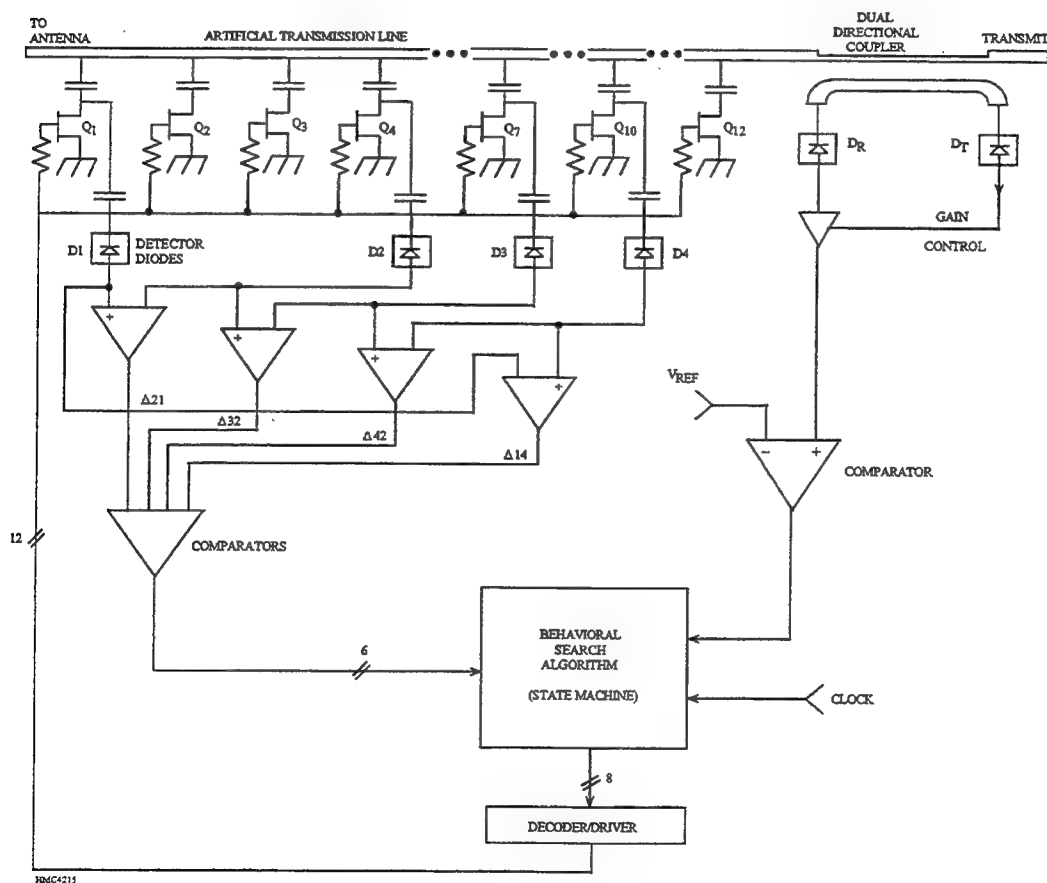
Figure 1. Block diagram of the dynamic matching network.

3.0   DISTRIBUTED MATCHING CIRCUIT CONSIDERATIONS

Since a distributed matching circuit consisting of FET switches varies only in discrete steps. It is necessary to determine minimum number of states required to perform the matching function to a load impedance of arbitrary phase angle. The design requirements are shown in Figure 2. The magnitude of the reflection coefficient of the load impedance ZL must lie within the boundary set by $|\Gamma_{LMAX}|$. The passive lossless matching circuit must then provide an impedance which guarantees that $|\Gamma_{S11'}|$ is less than $|\Gamma_M|$. A graphical representation of the problem is shown in Figure 3.
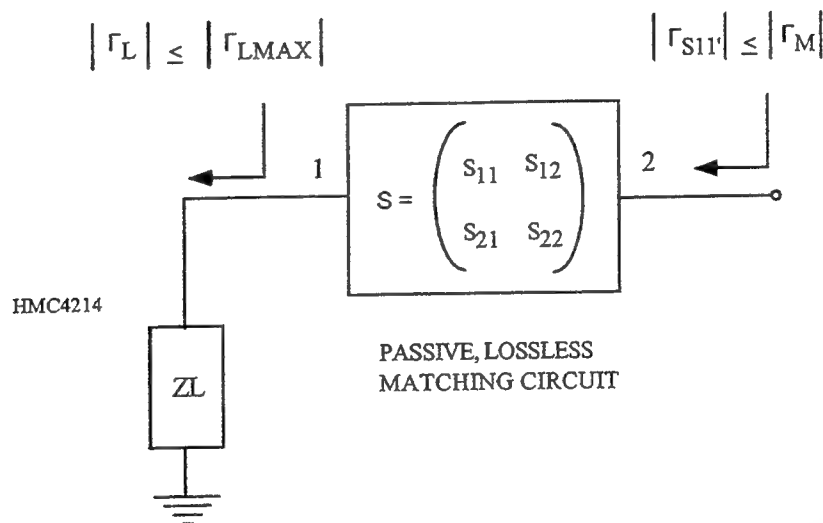
Figure 2.   Ideal matching circuit block diagram illustrating a linear, lossless matching network
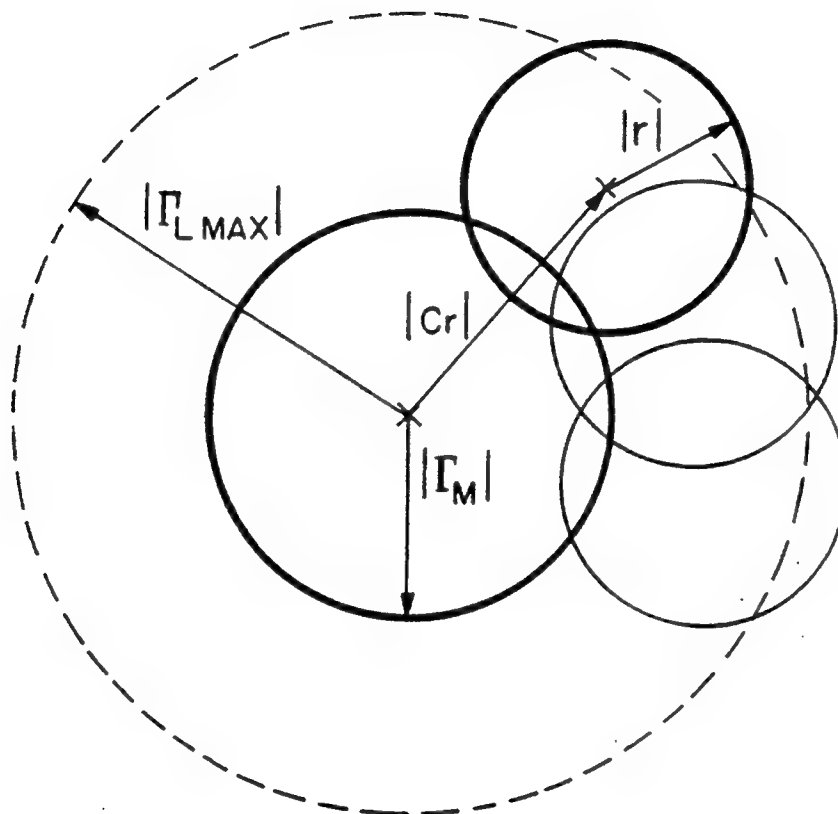
Figure 3.   Determination of minimum number of circuit states necessary for transforming any load with a 5:1 VSWR or less, into a net mismatch of 2:1 or less. For narrow-band applications, a minimum of 13 states is required.

It is required to match load VSWR's as high as 5:1 at any phase to VSWR's less than or equal to a 2:1. For narrow band operation these extremes are shown in the Smith Chart plane using VSWR circles of magnitude $|\Gamma_M|$, and $|\Gamma_{LMAX}|$. They represent the minimum and maximum boundaries of 2:1 and 5:1 respectively. At a given frequency we can define a circular area on the Smith Chart representative of the two port network such that any port one load impedance that lies within that area will be mapped within the circular area corresponding to a 2:1 VSWR. The center $|C_r|$, and radius $|r|$ are defined by equations 1 and 2. These quantities depend only on the minimum acceptable reflection (Γ_M) and the reflection coefficient (S_{11}) of the two port matching circuit. It can be shown that the ang C_r = -ang S_{11}.

$$|r| = \frac{|\Gamma_M|\ (1-|S_{11}|^2)}{1-|\Gamma_M\ S_{11}|^2} \tag{1}$$

$$|C_r| = \frac{|S_{11}|\ (1-|\Gamma_M|^2)}{1-|\Gamma_M\ S_{11}|^2} \tag{2}$$

The criteria for determining the minimum number of matching states required to cover the full range of load phase variation, is based on providing enough states (circles) so that the area bounded by the $|\Gamma_{LMAX}|$ circle is entirely covered by the state defined circles of radii $|r|$. To obtain complete coverage of the 2:1 and 5:1 VSWR circles, 12 circles are required each separated by 15°. A thirteenth state providing minimal reflection is also required for those matched loads that lie within the 2:1 VSWR circle.

The ideal matching circuit equivalent model, shown in Figure 4, consists of twelve unit length transmission line T-sections of length $\theta = 15°$ with switched in capacitors C_X. The distance θ is the angular difference between each state circle of Figure 3. For the thirteenth state, all switches are off and the circuit is a matched transmission line. The value of the capacitance C_X represents the capacitance change required at each switch location to transform the load mismatch to within the 2:1 VSWR circle. This can be calculated using equation 3. The corresponding simulated insertion loss and return loss for the ideal circuit into a load reflection $|\Gamma_L|=0.66$ are shown in Figure 5. The 0° trace corresponds to the correct switch location for a given load phase angle. The +15° and -15° traces correspond to adjacent switch locations. The distance between the points at which these adjacent switch traces intersect the 0° trace corresponds to the range of load phase angle that is covered by the 0° state.

413

$$C_X = \frac{2 \ |\Gamma_{LMAX}|}{2\pi f Z_0 \sqrt{1 - |\Gamma_{LMAX}|^2}} \tag{3}$$
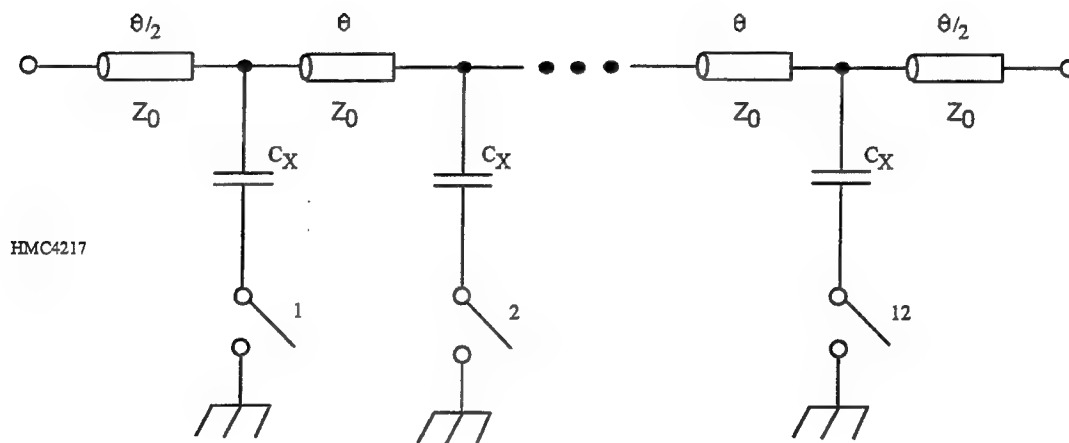
Figure 4. Ideal matching network consisting of sections of 50Ω transmission line sections with equally spaced shunt capacitances selectable with switches.
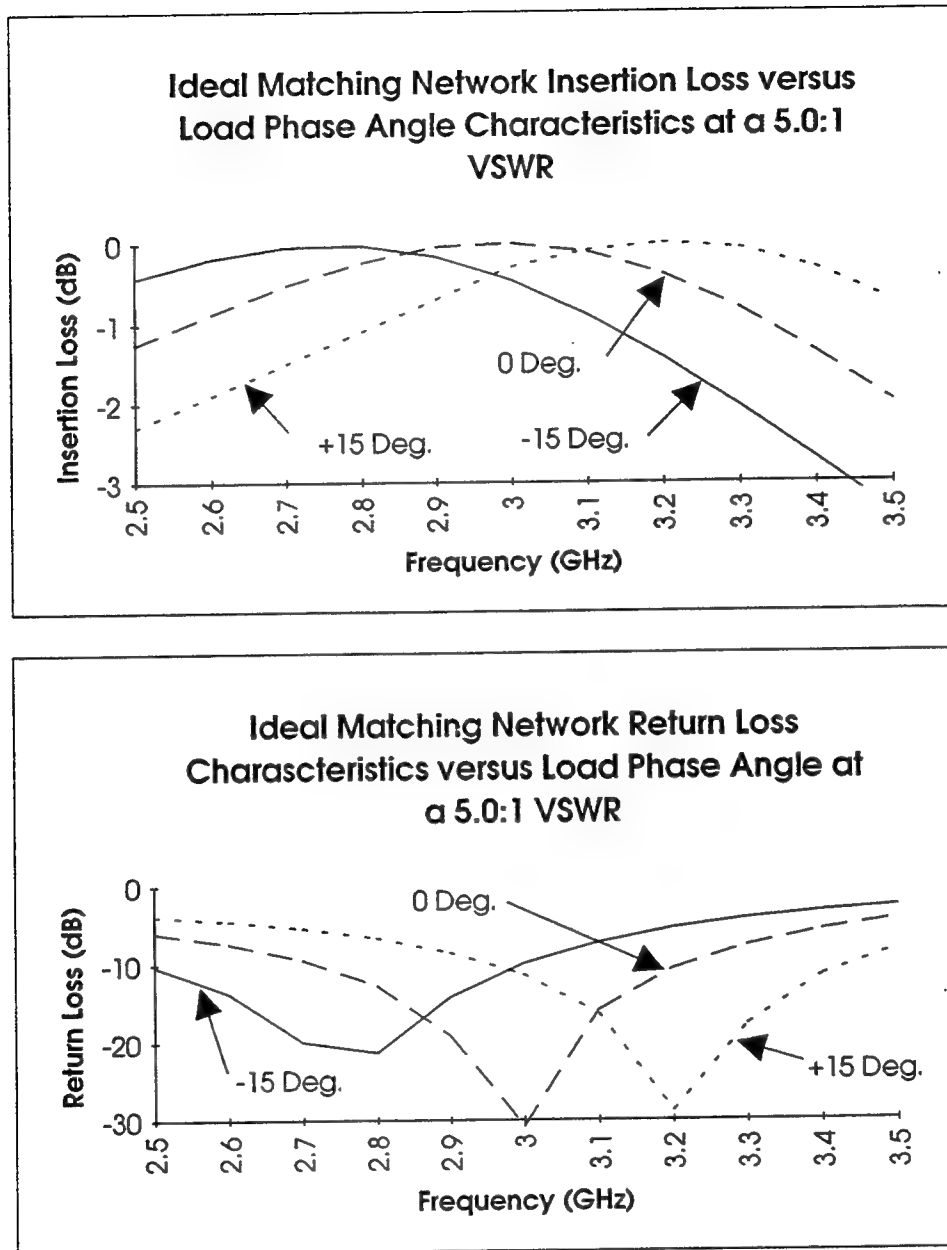
Figure 5. Ideal simulated insertion loss and return loss characteristics versus load phase angle for a 5:1 mismatch.

4.0   MATCHING CIRCUIT PHYSICAL REALIZATION

The approach used for determining the optimum matching network was to first select a switching device compatible with the maximum current and voltage requirements. The circuit shown in Figure 4 could then be modified to account for the switching device parasitics. The most critical parameter effecting the matching circuit passive (all off) performance, is the shunt capacitance of the switch, which lowers the impedance of the artificial transmission line. In the matched state (1 switch on), the FET ON resistance must be low to minimize resistive loss.

High power handling capability along with minimal control circuit power consumption requirements for the matching network led to choosing series stacked FET's as the switch elements.

The maximum peak voltage on the switch element occurs when the circuit is in the passive (all off) state, and the load mismatch is 5:1.   Under these conditions the maximum peak line voltage (V_L) is given by Equation 4.

$$V_L = (1 + |\Gamma_{LMAX}|)\sqrt{2P_{MAX}Z_0} \qquad (4)$$

Current in the shunt switch element however, peaks when the matching network is active and performing a match from a 5:1 to a 2:1. Under these conditions, the maximum peak current stress (I_{MAX}) can be calculated using Equation 5.

$$I_{MAX} = 2\pi f C_X (1 + |\Gamma_M|)\sqrt{2 P_{MAX} Z_0} \qquad (5)$$

Assuming a Γ_{LMAX} VSWR of 5:1 and a maximum signal power of 10 watts, the peak line voltage (V_L) determined by Equation 4 is 53 volts. The peak voltage across the FET switch is then reduced to approximately 44 volts due to C_X. From Equation 5, the peak current (I_{MAX}) is 1.5 A. The available series stacked FET MMIC's had been manufactured with a minimum drain to source breakdown voltage of 15 volts and saturation current (IDSS) of 350 mA/mm.

In order to handle a voltage swing of 44 volts, at least three FET's in series would be required. The only available three stack MMIC consisted of three 2.8 mm periphery devices. The peak current handling capability for the MMIC based on a maximum saturated current of 350 mA/mm, is 1 amp for the stack. To handle a peak ON state current of 1.5 A, two MMIC chips in parallel would be required.

417

The photograph for a three-stack FET MMIC chip is shown in Figure 6. The drain connection to the device are the three close proximity pads on the left side of the chip and the source connection are the three pads on the right hand side. The gate control connection is the upper right corner pad. The simplified chip schematic and approximate equivalent circuit for both the ON and OFF states are shown in Figure 7. If port 3 is isolated and port 2 is grounded, the OFF state equivalent circuit is 0.4 pF in parallel with a 700Ω resistor and the ON state equivalent is 1.0 pF in parallel with a 4.2Ω resistor.

The off state capacitance of the MMIC must now be absorbed into the matching circuit. This can be accomplished by increasing the impedance Z_0 of the transmission line segments. The ideal lumped element transmission line equivalents shown in Equations 6 and 7 can be used to analyze the effects of characteristic impedance change due to stray shunt capacitances.

$$\theta = 2\pi f \sqrt{LC} \qquad\qquad (6)$$

$$Z_0 = \sqrt{\frac{L}{C}} \qquad\qquad (7)$$
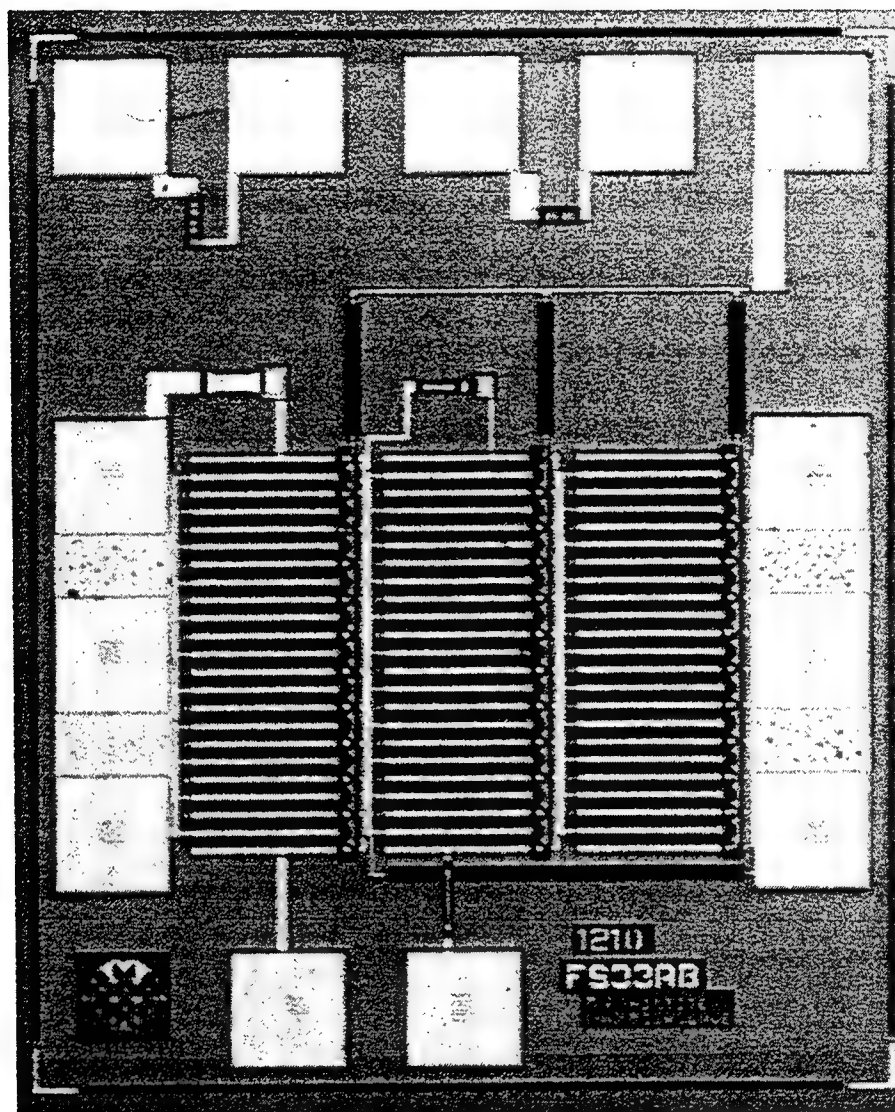
Figure 6.   Photo of a series three stack MFET MMIC. Each FET is 2.86 mm in
total periphery. On chip compensation capacitors are used to equalize
the switching delays during high power switching applications, thus
preventing individual device burnout.

419
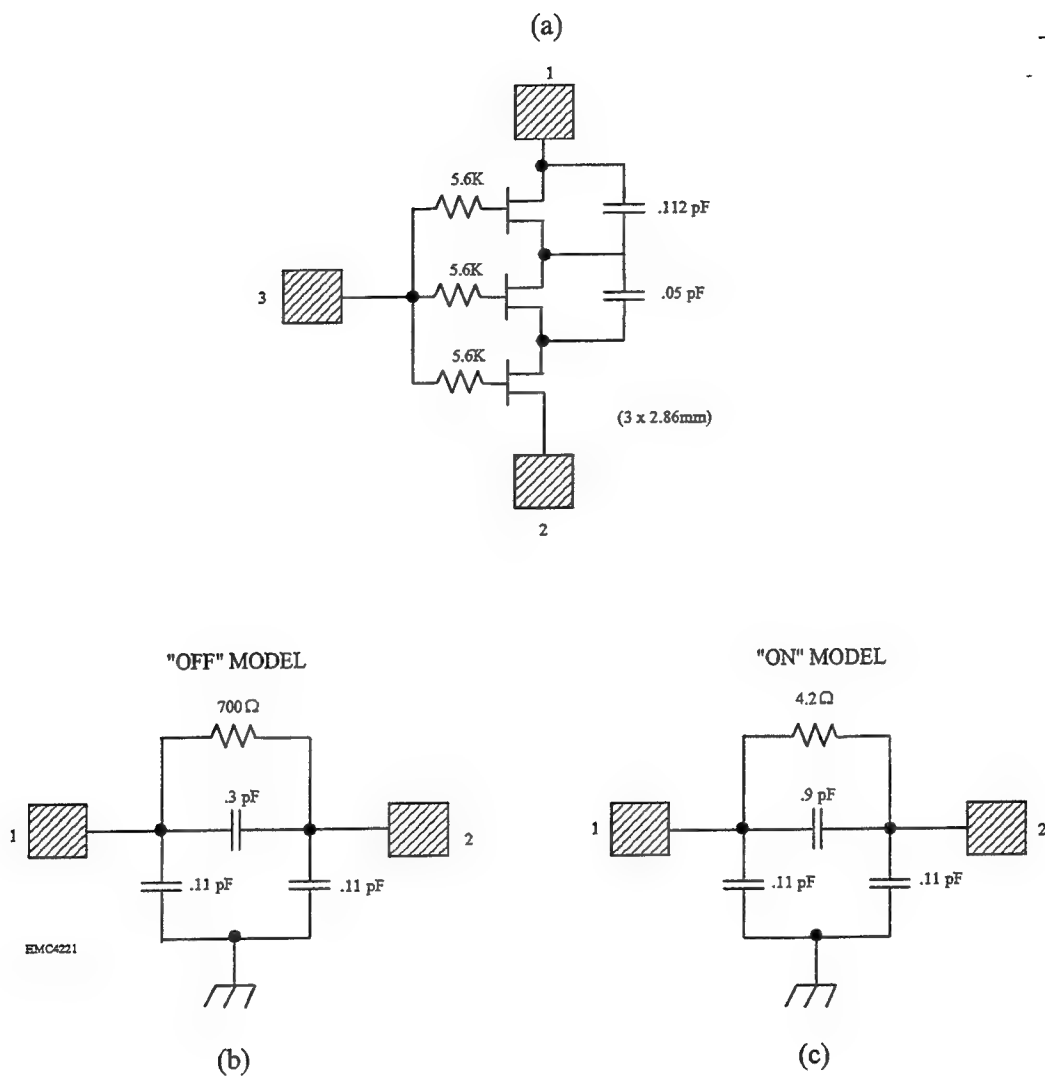
Figure 7.   (a) Simplified three stack MMIC schematic. (b and c) Approximate

equivalent circuits for both the off and on states of the MMIC with

port 3 isolated.

Using these equations, it can be shown that if two MMIC's are placed in parallel at each switch location, the off state capacitance is too high to maintain a 50Ω artificial transmission line with an electrical length of $15°$. To overcome this problem, a thirteenth switch was added and devices were turned on in pairs. The transmission line lengths then became $13.8°$ for 13 switches instead of $15°$ for 12 switches. This allowed the current handling benefits of paralleling two devices, without the penalty of twice the off state shunt capacitance at each node. Furthermore, by turning on the devices in pairs, the on state resistance is effectively reduced by a factor of two, thus improving the networks insertion loss in the matched states.

The final step in the matching network design was to replace the inductors with their transmission line equivalents and modify the previously calculated value of C_x, considering the parasitic capacitances associated with the transmission line inductors and the MMIC switches. Optimization was performed for insertion loss and VSWR in both the active and passive states using accurate linear models for the MMIC devices.

5.0   CONTROL ALGORITHM

A method of determining the correct switch location to compensate for any given load mismatch in the range of 2:1 to 5:1 at any phase was needed. Various

algorithms for controlling the matching network to minimize VSWR, such as look up tables and random searches were investigated. The most appropriate algorithm for application in T/R modules would be that which minimized VSWR settling time and eliminated the possibility of creating a greater mismatch during the search process. This involved calculating the optimum switch location for the standing wave pattern present on the line for any given load reflection.

The slope detection algorithm provides an initial estimate of the optimum state before the search begins. The optimum stub location will then fall within a group of a few devices. An initial estimate is obtained by measuring the slope of the peak amplitude vs. distance along the transmission line.

This concept is illustrated in Figure 8. Γ_θ represents the load reflection as viewed from the transmission line. The angle θ is the distance from the load to the point in which the reflection coefficient falls on the unity conductance circle. It is at this point that the circuit can be matched with a simple shunt capacitor. As this point on the transmission line moves away from the load, the angle θ rotates clockwise. The $1+\Gamma_\theta$ phaser is tangent to the circle defined by Γ_θ. This is the point in which the rate of change of slope is maximum, thus is the location in which the capacitor should be applied. Because the value of capacitor is not optimum for all possible loads, uncertainty is introduced and the selected switch position could be off by as much as an adjacent location.
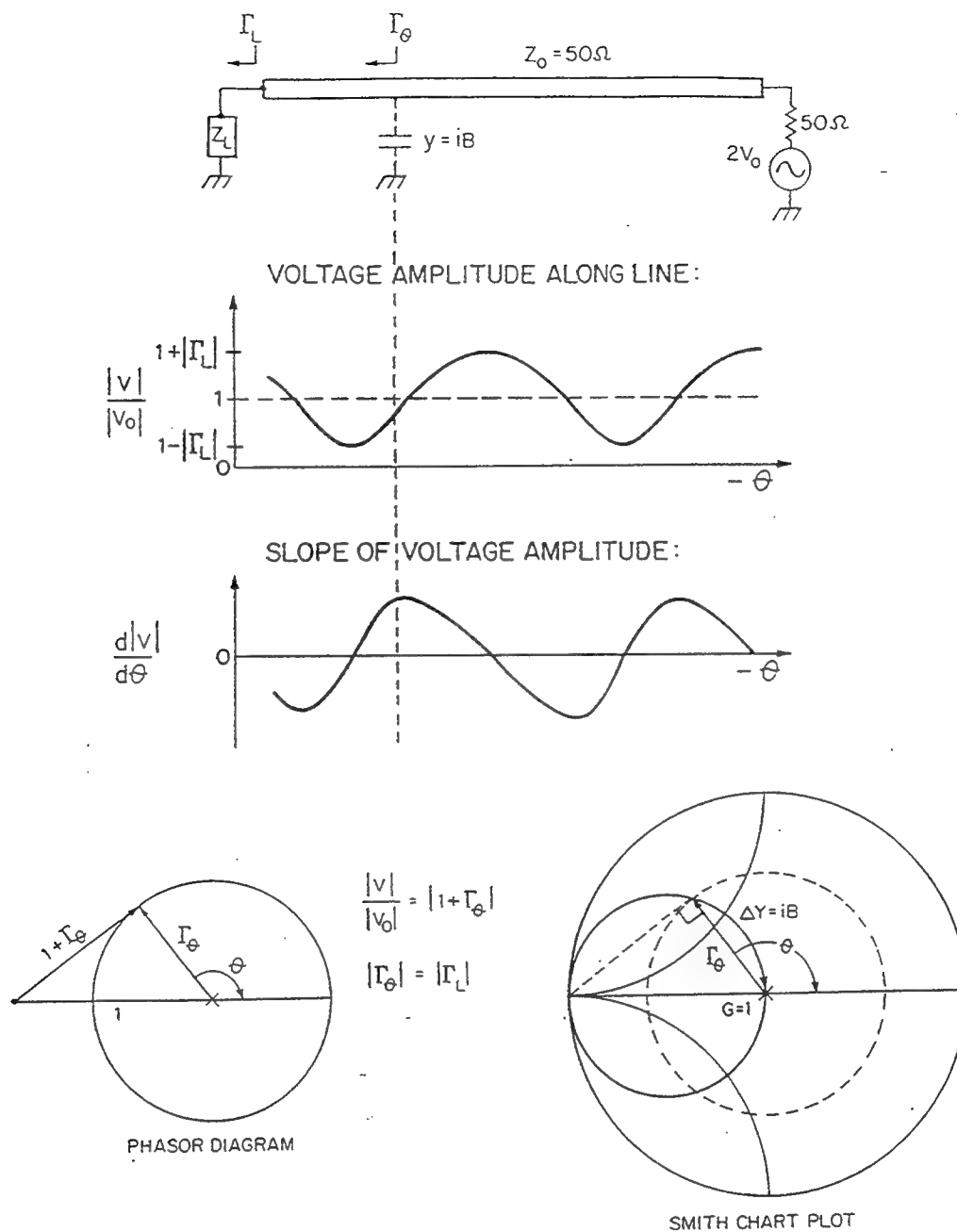
Figure 8.   Principle of slope detection for the adaptive matching control circuit. The matching element is applied at the point of maximum slope amplitude.

6.0   FABRICATION

The microwave module was fabricated in a copper housing 1.6 inches square by 0.5 inches tall. The artificial transmission line was implemented on a duroid substrate to minimize insertion loss. The dual directional coupler however was fabricated on an alumina substrate due to size and tolerance constraints. The three-stack MMIC chips were mounted on thermcon heat spreaders recessed into the housing on both sides of the artificial transmission line. Transmission line connections to the switch MMIC's and standing wave detector diodes were made using high Q chip capacitors. A quarter wavelength short circuit transmission line filter had to be added between the dual directional coupler and the artificial transmission line to attenuate the second harmonic of the operating frequency at the incident and reflected detector diodes. The detectors were implemented using EFET's configured as diodes.

All detector amplifiers, comparators, digital circuitry, and FET switch drivers were contained on an external PC board using a combination of through hole and surface mount chips. The digital state machine was implemented within a programmable array logic chip. Photographs of the microwave module and external PC board are shown in Figures 9 and 10.
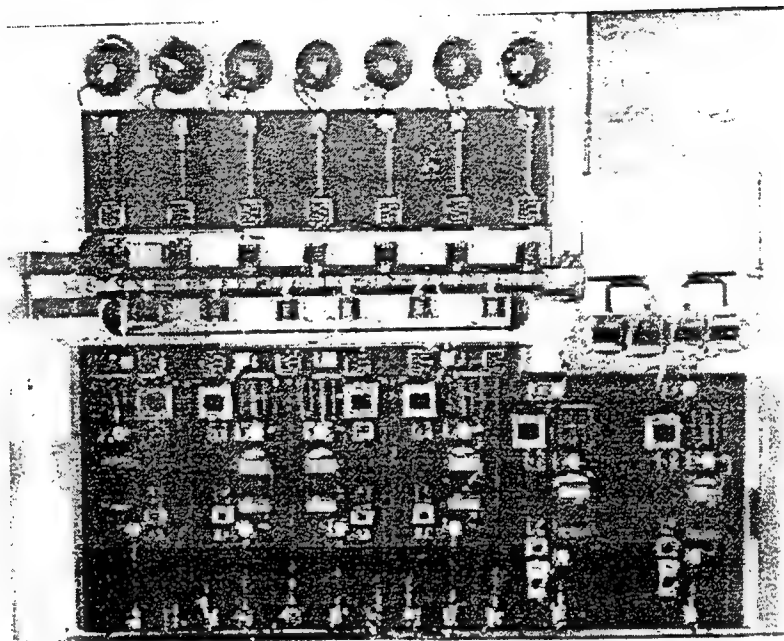
Figure 9.   Photograph of the adaptive matching network microwave module.
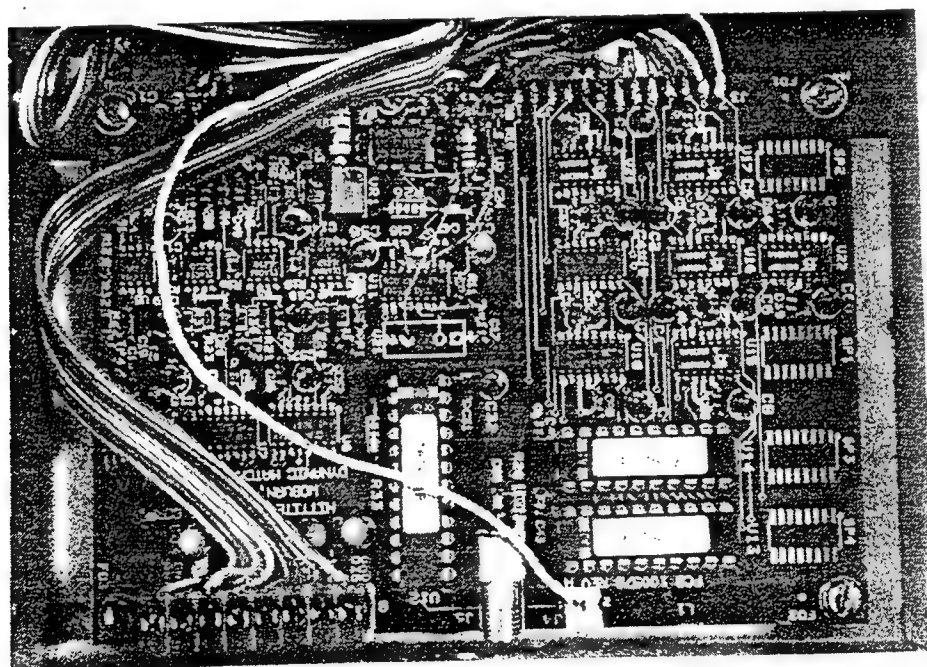
Figure 10.   Photograph of the control circuit PC board.

7.0 TEST RESULTS

Final circuit testing was performed at a frequency of 3 GHz and a power level of 2 watts. Load mismatches of 2:1, 3:1, 4:1, and 5:1 were rotated in phase in 11 increments so as to cover the entire smith chart. At a VSWR of 2:1, the circuit sensed that the mismatch was not severe and in 10 out of 11 phase positions remained in it's passive (all off) state. In one phase position however, the VSWR detector sensed a mismatch above the minimum threshold and re-matched to better than a 1.5:1. All other mismatches greater than a 2:1 were properly sensed and re-tuned to better than a 2:1. Plots of the measured VSWR and Insertion loss performance is shown in Figure 11. Note that half of the Smith chart can be covered with a corresponding insertion loss less than 2.4 dB.

Typical and worst case settling time plots for the adaptive matching network are shown in Figures 12a and 12b. The upper trace is the active low reset pulse. All measurements are referenced to the rising edge of this pulse. The second trace is the external clock. The third trace represents the difference between the incident and reflected powers. The bottom trace is the output of the VSWR threshold comparator, and if low, corresponds to a match condition better than a 2:1 VSWR. The measurements were performed at 5 MHz which represents the maximum clock speed in which the internal analog and digital signals have time to properly settle. The
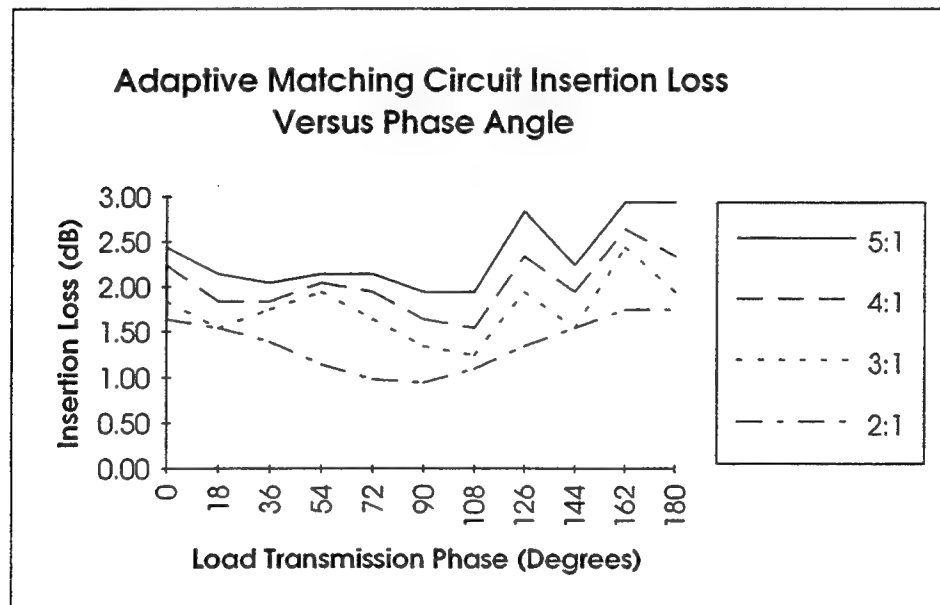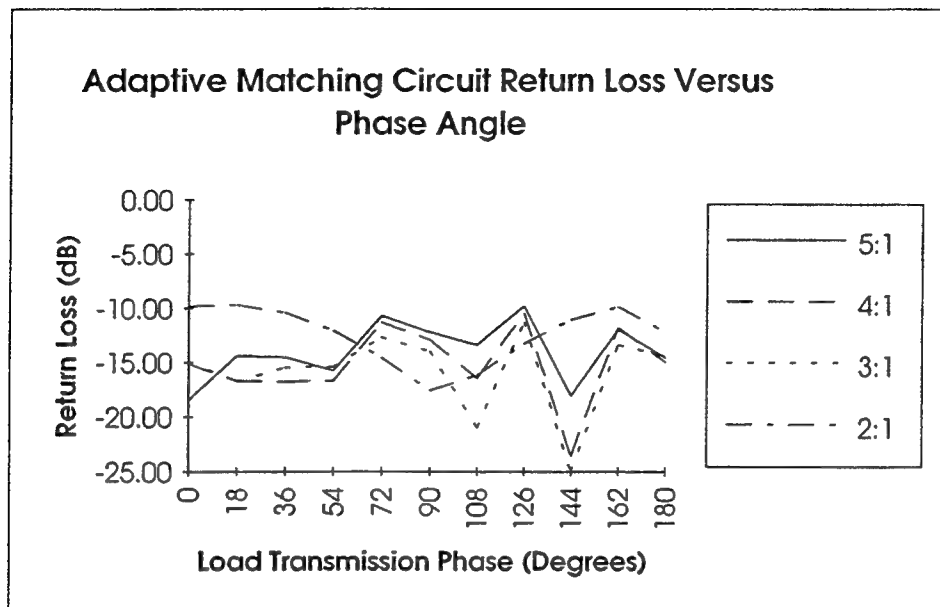
Figure 11.   Final VSWR and insertion loss performance of the adaptive matching

network at various phase and mismatch conditions.

427
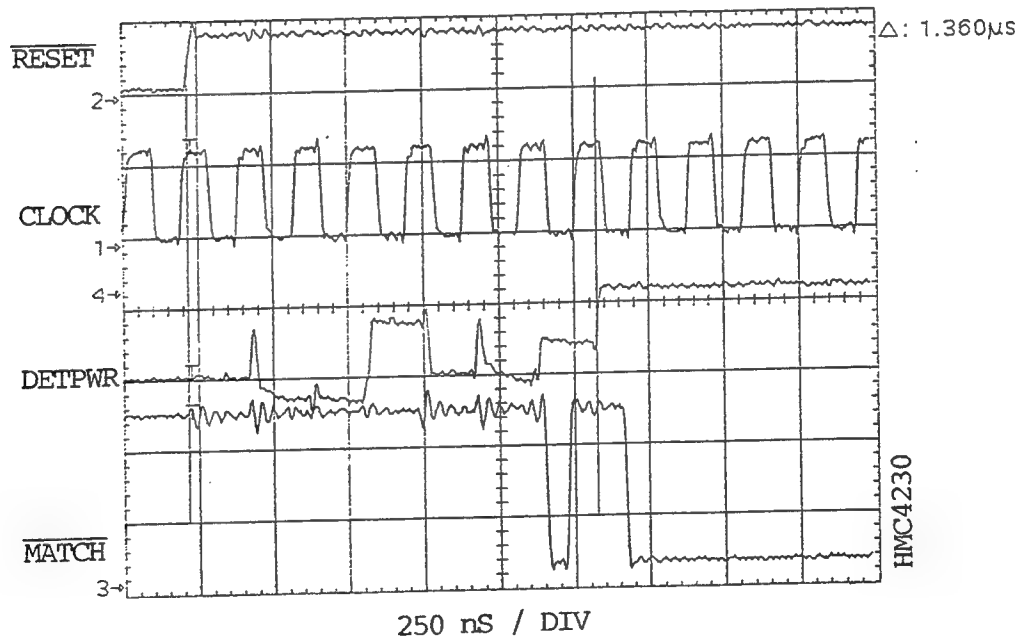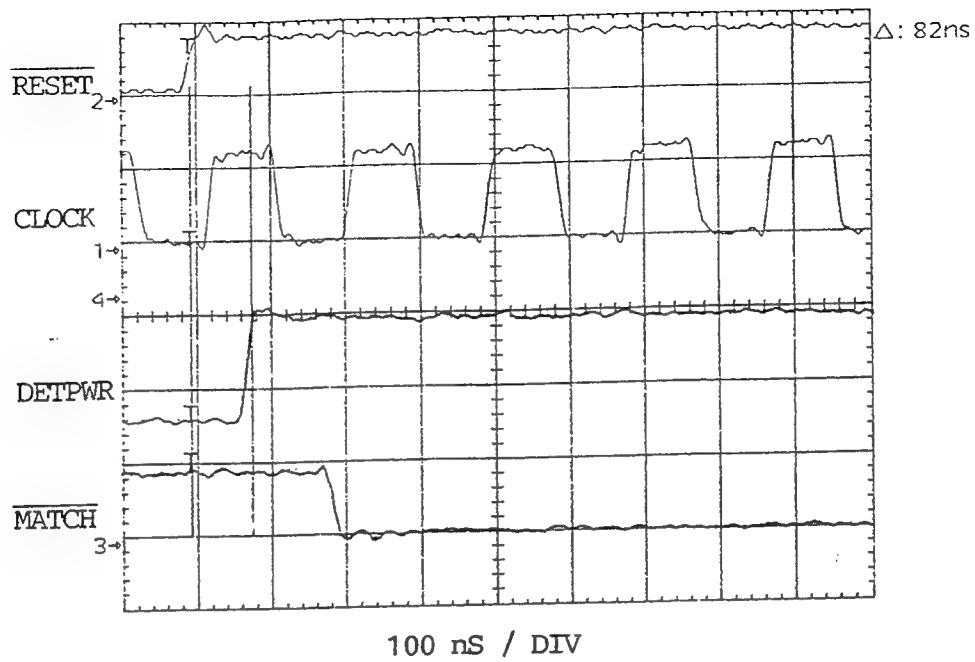
Figure 12. Typical and worst case settling time measurements for the adaptive
matching network.

428

fastest settling time observed was 82 nS which corresponds to less than one clock cycle. The longest settling time was 1.36 uS which corresponds to seven clock cycles.

8.0   CONCLUSION

An MIC version of an adaptive distributed matching circuit which can transform load mismatches up to 5:1, at any phase angle, to better than a 2:1, has been designed and fabricated. The design was centered at 3 GHz and all testing was performed at a power level of 2 Watts. Average insertion loss performance over all states and load VSWR's is < 2.0 dB, with slightly higher loss at 5:1 mismatches in the half of the Smith Chart furthest away from the load. The average settling time is less than 500 nS, and the worst case is less than 1.4 uS.

Insertion loss performance was slightly higher than that demonstrated previously [2]. This was due to a limited selection of series stacked MMIC switches. If the periphery of the stacked fets were increased, the drain to source resistance in the ON state could be reduced, resulting in an insertion loss less than 2 dB.

There was a significant difference between the best and worst case settling times. This problem can be resolved in the future by slightly increasing the size of the programmable array logic chip, as the existing chip was fully utilized. The present algorithm looks for the maximum positive slope of the standing wave pattern and

429

determines which group of switch devices to choose as the starting point for the search. If the algorithm were modified to input the location of the next most positive slope on the transmission line, then the search for the optimum match switch element would proceed in the proper direction. This would minimize the total number of clock cycles to arrive at the optimum match state.

ACKNOWLEDGEMENTS

REFERENCES

[1]   R. Chu and Lee, "Radiation impedance of a dipole printed on periodic dielectric slabs protruding over a ground plane in an infinite phased array," IEEE Trans. Antennas Propagat., vol. AP-35, pp 13-25, Jan 1987.

[2]   A. W. Jacomb-Hood, "A Novel Aperture Isolation Circuit for Use in Phased Array Systems" IEEE Trans MTT, vol. 38, No. 12, Dec 1990.

[3]   M.B. Shifrin, P. Katzin and Y. Ayasli, "Monolithic FET structures for high-power control component applications", IEEE Trans. MTT vol. 37, pp.2134-2141, December, 1989.

[4]   P. Katzin, B.E. Bedard, M. Shifrin and Y. Ayasli, "High-speed 100+ Watt RF Switches using GaAs MMICs", IEEE Trans. MTT vol. 40, pp 1989-1996, November, 1992.

An Active Monolithic Programmable Directional Coupler for Phased-Array Radars

Wes Grammer, Mitchell Shifrin, Christopher Lyons, Peter Katzin

HITTITE MICROWAVE CORPORATION

ABSTRACT

An active S-band microwave directional coupler with programmable coupling over a 30 dB range has been demonstrated. The coupler circuit uses a novel planar transformer design, integrated with a variable stepped-gain amplifier on a single MMIC, suitable for integration into a T/R module. In addition, a phase-compensated variable attenuator MMIC was demonstrated, with better than 3° RMS phase variation.

1.0 INTRODUCTION

Directional couplers are one of the principal elements in beam-forming networks for phased-array antennas, in applications such as monopulse radar. In systems that use an array of monolithic T/R modules, a coupler in MMIC form could be integrated in the module, saving space and lowering overall cost. Figure 1 illustrates the functions of the coupler in a monopulse system. In transmit mode, the coupler must pass the exciter signal with minimum loss and phase distortion. In receive mode, the coupled port amplitude must be controllable over a broad range, without affecting other coupler parameters such as directivity and isolation. The

432
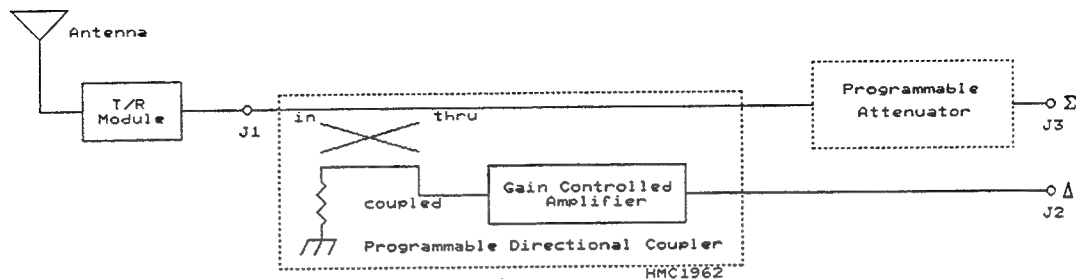
Figure 1.   System block diagram, programmable coupler and attenuator.

variable attenuator shown in the sum channel is another element in the beam forming network, and must have minimum phase change vs. attenuation, and good linearity. This paper presents the design of both components as GaAs MMICs. Also, measured data from each is presented.

The programmable directional coupler MMIC consists of two elements: a lumped-element, broadband coupler with a novel planar transformer layout, and a variable-gain amplifier, adjustable over a ±15 dB range in 30 linear steps. The variable attenuator MMIC is digitally programmable in 1 dB increments, with 5 bits (32 dB) of range. Both MMICs include compensation networks to minimize the insertion phase variation over the range of settings.

2.0 LUMPED-ELEMENT COUPLER

GaAs MMIC power dividers and couplers have been demonstrated using lumped-element or distributed techniques [1-4]. However, these designs are prohibitively large at S-band, and do not possess the low thru insertion loss and good directivity needed in a programmable coupler application. An alternative approach is a lumped-element coupler made with two identical, high turns-ratio transformers, as shown in Figure 2. Couplers of this type are widely used at low RF frequencies. The capacitors C_1 and C_2 compensate for the interwinding parasitic capacitance of the transformers, greatly improving the coupler directivity. The L-C network in the thru path linearizes the phase response and improves port match over the band of
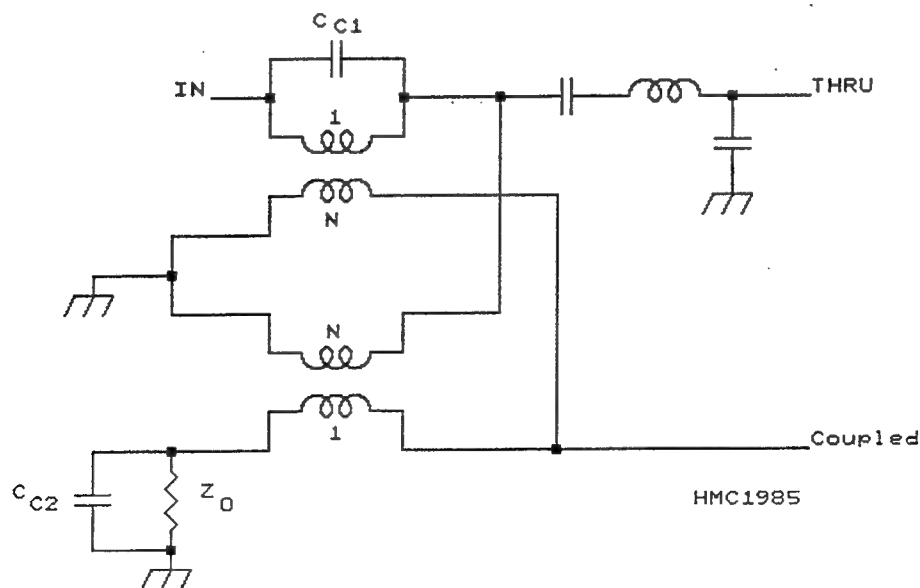
Figure 2.   Circuit Diagram, Lumped-Element MMIC Coupler.

interest. To realize the transformers in planar MMIC form, the layout in Figure 3 was devised. The secondary winding is a single spiral-wound coil, interwoven with four parallel-connected loops that form the primary. Both windings are airbridged over most of their lengths to minimize capacitive coupling and loss. The design yields a coupling factor K~0.6, determined from a calculation of the self and mutual inductance of the coils using quasi-static approximations. Total size of the coupler test cell is 1000 x 1300 μm^2.
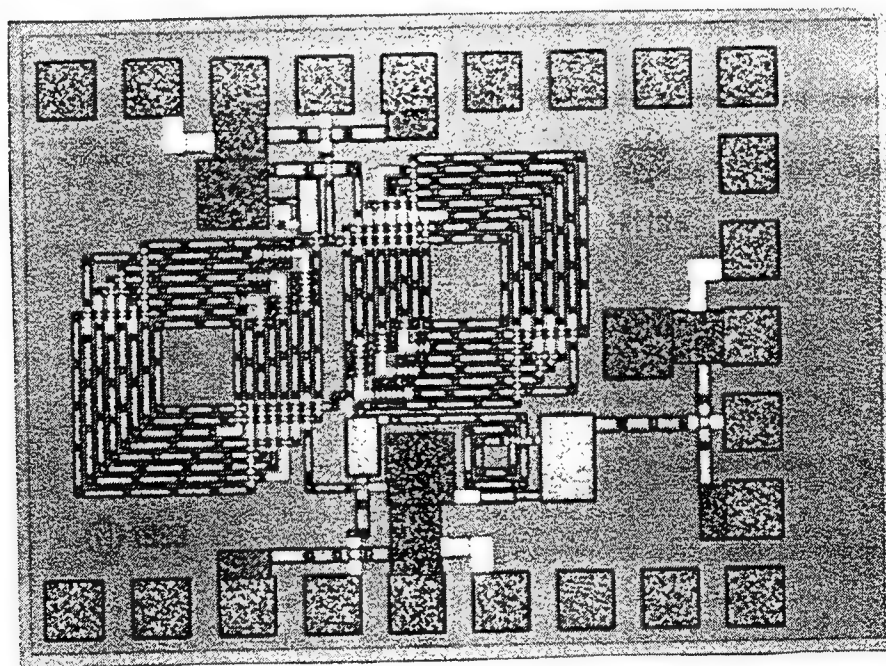
Figure 3.   Layout of coupler test cell.

3.0   VARIABLE-GAIN AMPLIFIER

Figure 4 shows the circuit diagram of the amplifier design, including bias networks. A two-stage common-source configuration is used, with lossy interstage and port matching networks to give a flat, broadband gain response and low VSWR. Both stages are biased at I_{DSS}, to eliminate the need for gate bias. The first stage consists of five parallel dual-gate MESFETs, with peripheries scaled in ascending binary order. The second gate functions as a control bit, selectively enabling or disabling a given device. Thus, the composite FET has a transconductance g_m (and hence a gain response) linearly proportional to the binary control inputs. To compensate for a slight shift in the insertion phase response versus gain, a small amount of shunt capacitance is added progressively at the lower gain settings. The second stage also uses a dual-gate FET, in this case to improve isolation of the output load from the variable-gain stage.

The layout of the variable-gain amplifier, configured as an individual test cell, is shown in Figure 5. A 1μm Triquint process with dielectric vias was used for device fabrication. The cell measures 1.0 x 1.3 mm, from the outside edge of the bonding pads.
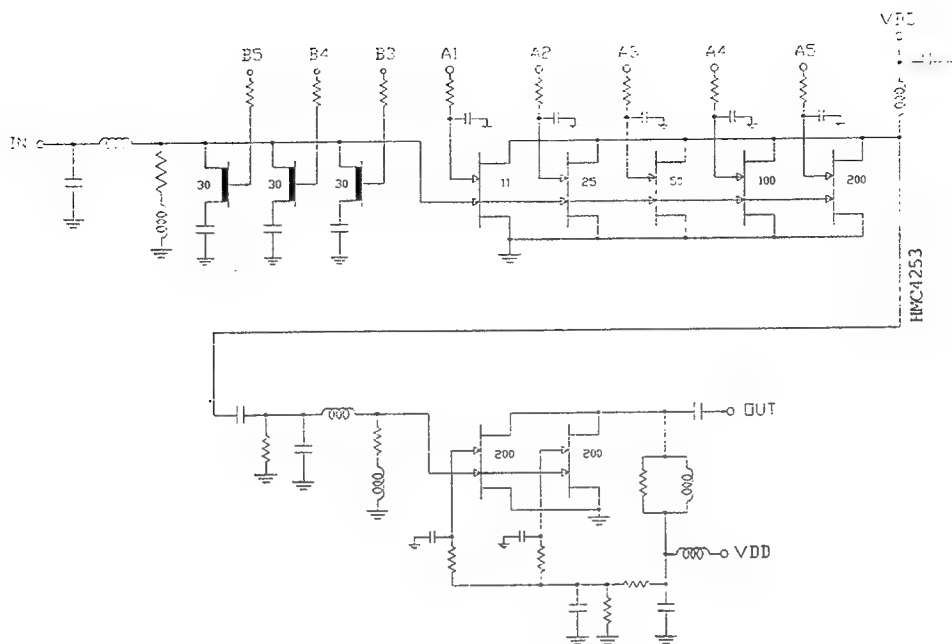
Figure 4.   Circuit diagram, variable-gain amplifier.
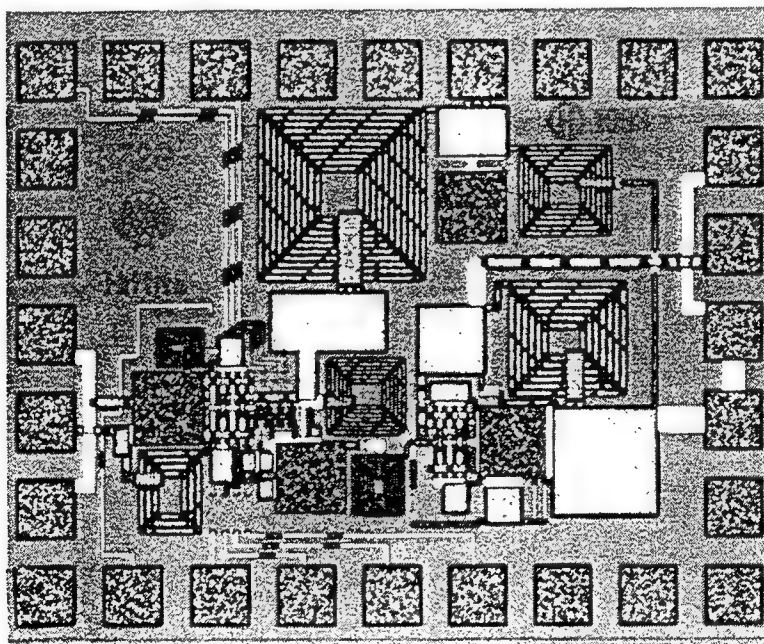
Figure 5.   Layout of amplifier MMIC test cell.

437

4.0 INTEGRATED COUPLER/AMPLIFIER

The integrated programmable coupler layout is shown in Figure 6. A small capacitance was added at the amplifier input, to compensate for the inductance of the line connecting to the lumped-element coupler. Also, some of the bias network components were relocated to make the design more compact. Otherwise, the amplifier layout is identical to the test cell described earlier. Overall, the cell measures 1.75 x 1.45 mm, including the bonding pads.

The RF input and output (thru) ports are on the top and left sides, respectively. The coupled port output is on the right side, and the bottom row of pads are for setting the gain of the coupled signal, and for drain bias. Grounded pads are provided on either side of the RF inputs to allow wafer probe testing.
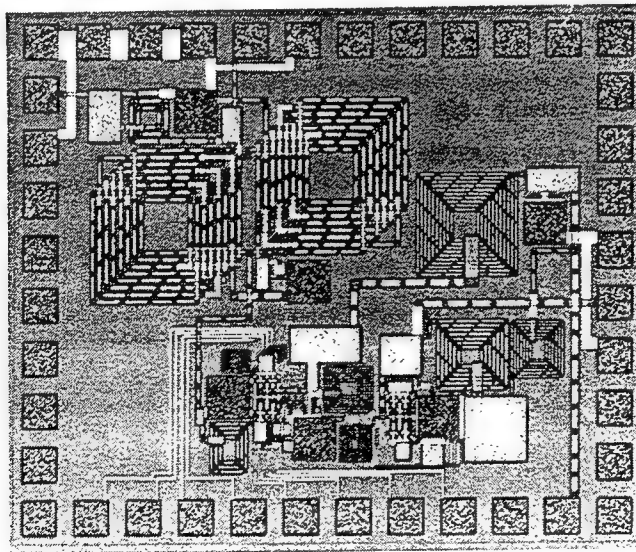
Figure 6.   Layout of programmable directional coupler MMIC.

5.0 VARIABLE ATTENUATOR

Figure 7 shows the circuit diagram of the passive 5-bit digital attenuator. The topology uses cascaded sections of 4, 8, 1, 2 and 16 dB, fabricated with MESFET switches and fixed resistors. In the 1 and 2 dB sections, a shunt resistor is switched in, to give the appropriate attenuation. Degradation in port match is not severe enough to require a more elaborate circuit, and the OFF state thru loss is very low. The 4 and 8 dB sections are elementary pi-attenuators, and use a pair of shunt FETs with a single series FET, switched in complementary fashion. The 16 dB stage is merely a cascade of two 8 dB stages; prior analysis had revealed this configuration to outperform a single-stage design. The series inductors and shunt capacitors included within the attenuator sections were added to equalize the insertion phase in the ON and OFF states. Additional inductance was included between stages to improve the OFF-state match.

The attenuator circuit was realized in MMIC form using a 1-μm Triquint process, with substrate vias. A low-inductance ground path was found to be very important in reducing the amount of added reactance needed for phase compensation. Thin-film NiCr resistors are used for all critical values, to insure accuracy, stability, and good yields. The layout of the attenuator MMIC is shown in Figure 8.
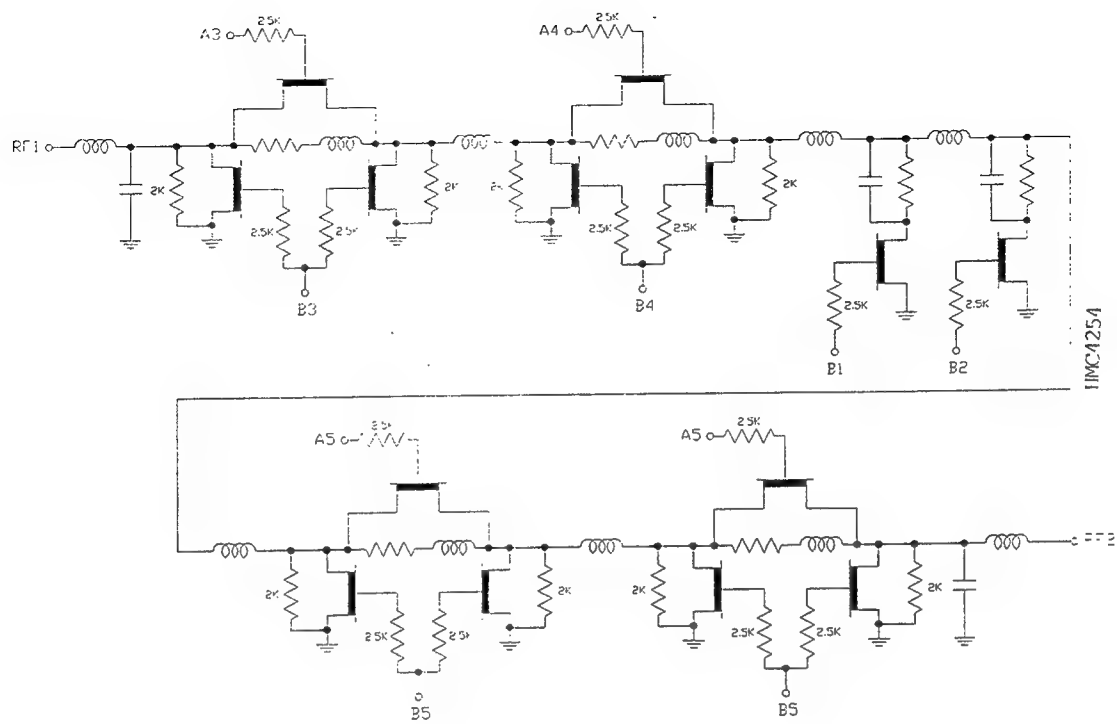
Figure 7.   Circuit diagram, variable attenuator.
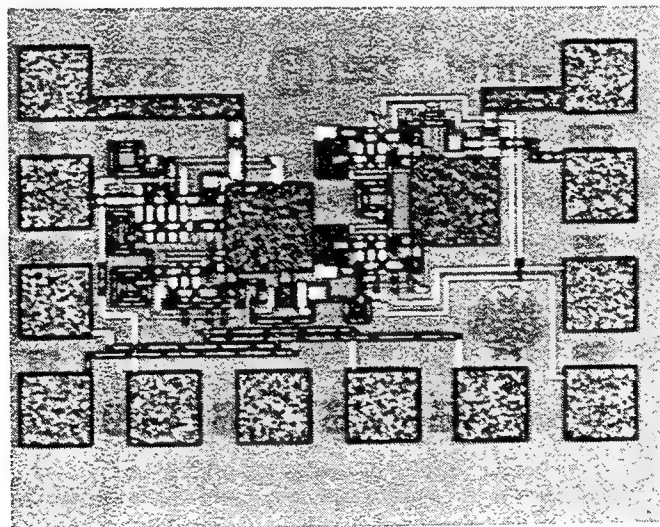
Figure 8.   Layout of variable attenuator MMIC

6.0   MEASUREMENTS

The variable attenuator and programmable coupler MMICs were measured on-wafer using an 8510B network analyzer, from 2.5-5.0 GHz. In addition, a test fixture containing the MMICs in the configuration of Figure 1 (minus the T/R module) was assembled and measured.

Figure 9 shows the measured thru loss of the attenuator for all 32 states, and a plot of the insertion phase, both normalized to the minimum loss ("0 dB") state. The amplitude response is monotonic and flat, with a slight slope at the higher-loss states. Peak-to-peak phase change (all states) is ~7°, worst-case, with a computed maximum RMS phase error of less than 3°. Finally, the minimum loss of the attenuator averages around 6.5 dB, with about a ±0.2 dB variation across the band.

The thru loss, coupling, and isolation of the lumped-element coupler test cell are plotted in Figure 10. The thru loss is low, and the coupling broad and flat at -15dB. However, the directivity is much lower than that predicted in our simulations (> 20 dB). The exact reason for this is not known at present, but maybe due to unwanted coupling between the two transformers, an effect that was not modeled.

Lastly, the measured coupled port insertion loss of the programmable coupler chip is shown in Figure 11(a), for all 32 gain states of the amplifier. Because the scaling of the voltage gain with setting is linear, the responses appear exponential on

a dB scale. With a predicted gain of ±15 dB, the results are reasonably close to that predicted in the design simulations. In Figure 11(b), the phase deviation, (normalized to the highest gain state) shows a fairly large peak error for the lowest 3-4 states, reducing the useable dynamic range by about half. Further improvement in the model used for the segmented dual-gate device in the first amplifier stage may improve the phase matching at lower gain settings.
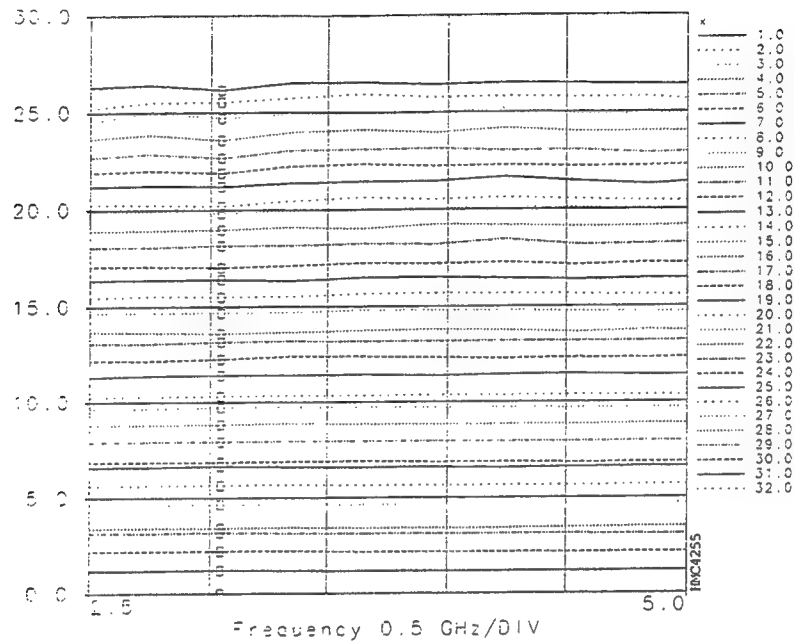
a) Magnitude (dB), normalized
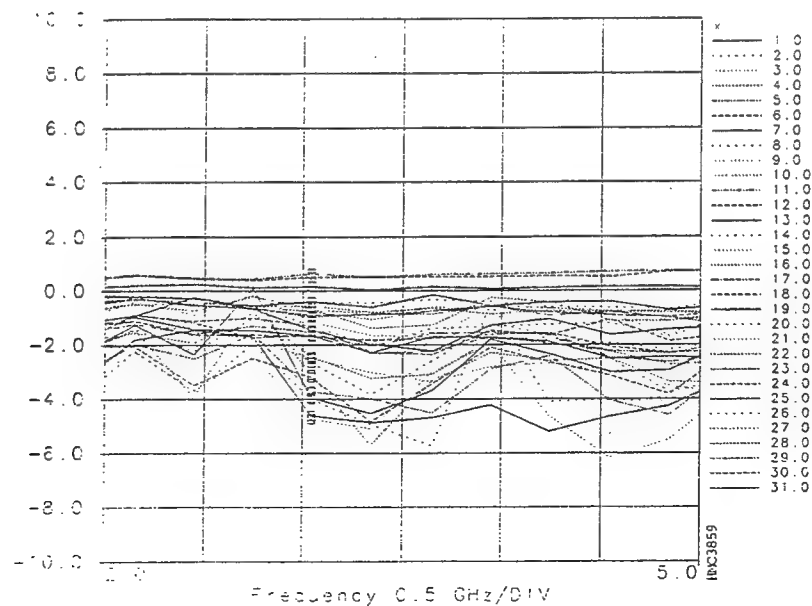
b) Phase (deg.), normalized

Figure 9.   Measured response, variable attenuator MMIC.
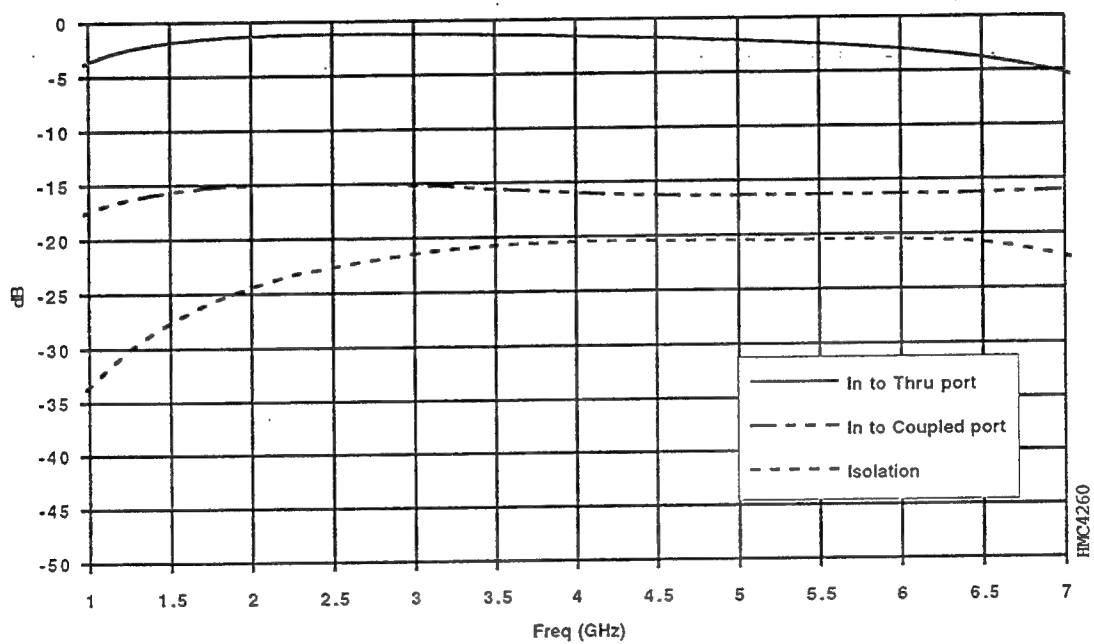
Figure 10.   MMIC coupler response.
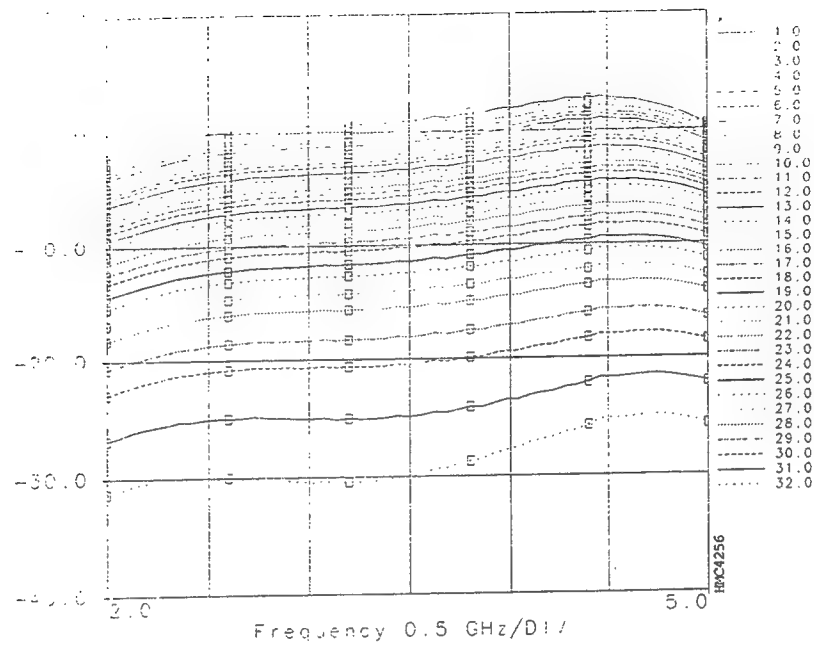
a) Magnitude (dB)
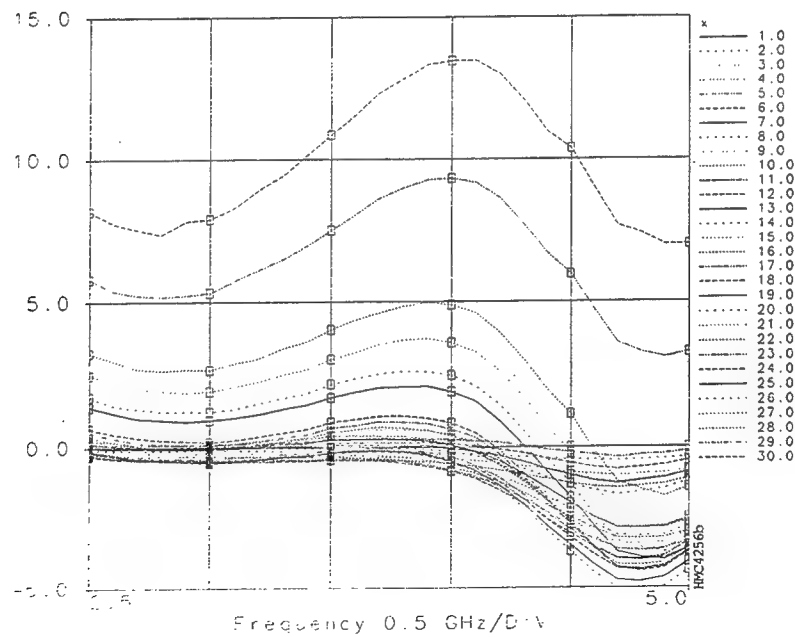
b) Phase (deg.) normalized; two lowest-gain states excluded

Figure 11.   Measured response, programmable directional coupler.

7.0 CONCLUSION

A programmable directional coupler MMIC, suitable for use in phased-array antenna beam-forming networks, has been designed and tested. The coupler has low insertion loss, good directivity and linearity at a broad range of coupling values.

REFERENCES

[1]   R.E. Neidert, S.C. Binari, "mm-Wave Passive Components for Monolithic Circuits", MW Journal, April 1984, pp. 103-120.

[2]   W. Beckwith and J. Staudinger, "Wide Bandwidth Monolithic Power Dividers", Microwave Journal, Feb. 1989, pp. 150-160.

[3]   J. Putnam and R. Puente, "A Monolithic Image-Rejection Mixer on GaAs Using Lumped Elements", Microwave Journal, November 1987, pp. 107-116.

[4]   G.E. Brehm and R.L. Lehmann, "Monolithic GaAs Lange Coupler at X-band", IEEE Trans. on ED, Vol. ED-28, No. 2, February 1981, pp. 217-218.

FROM DC(ØHZ) TO DAYLIGHT(10^{14}HZ);DEVICES,USES,AND
SIMILARITIES

George J.Monser

GJM Consulting Engineer
8516 W.Lake Mead Blvd.,Suite 199
Las Vegas,NV 89128:(702)656 2319

Abstract:Commonly used techniques in optical and
antenna design are shown and related.Hall devices
(capable of non-reciprocal d-c reception)are comp-
ared to loops and whips for low frequency use.

1.0 Introduction

Hall devices [1] represent the lowest frequency known
responding to incident magnetic flux density,so that d-
c reception is possible;moreover the device is non-re-
ciprocal and cannot be used for radiation.In practice,
Hall sensors have been used from 10 cps to 15Khz with
nearly frequency-independent sensitivity.They offer an
alternative to loops and whips which are both frequency
sensitive and narrow-band when tuned.

Classically,antennas have been designed for use from at
least 1 Mhz to beyond 60 Ghz.Over this range frequency
independent antennas have found wide use.Typically,they
work over multioctave bands.

Over the microwave band (1 to 60 Ghz) reflector-fed an-
tennas,gridded-reflectors,lens-antennas,and notches are

extensively used.

In the optical domain [2] ,wavelengths from 400 to 700 milli-microns dichroic reflectors have been used to split energy into two sub-bands.

This paper describes some of these devices and their similarities.

2.0 Hall Devices

Figure 1 shows a Hall sensor with enhancement flux collectors.The output voltage and flux are related by,

$$V=hB.....................................(1)$$

Where:

h=RI/t, is the transfer constant,R is the Hall coefficient,I is the bias current in the sample, and t is the thickness of the sample.

B= incident flux density on the sample.

The relationship between electric field intensity,E, Hall voltage,V,and internal noise voltage are found. Equating V to the noise voltage allows determining the minimum field voltage,E.

Table 1 shows the results compared to a short whip and loop for the same bandwidth.Notice that the loop and whip are much more frequency sensitive than the Hall device.The Hall device source impedance is

resistive and hence better suited for broadband uses.

3.0 Microwave Lenses

The Gent [3] or Rotman [4] lens is offered as an ex-example (Figure 2).

Four design features exist;

1. Shape of input surface

2. Shape of the output surface

3. Positions of the pickup elements

4. Lengths of the transmission lines, lens to array

As a result three perfect focus points can be allocated along the input surface. The surface facing the array can be arbitrarily selected to match linear or curved array geometries.

Beampointing is accomplished by movement along the focal-arc surface. The design yields a true time-delay system so that beams are fixed in location (not frequency dependent).

The significant point here is that geometric optics was used to develop the design.

4.0 Polarization Grids

Figure 3 shows one type grid. Here, when the incident field is aligned with the grids reflection will occurr. Orthogonal fields will pass through except for some

backscatter.For close spacing (0.1 or less in wave-
lengths) nearly perfect reflection will result.
These properties have led to applications in sub-re-
flectors as band-splitters,polarization-splitters
and suppressors.

5.0 Dichroic Beam-Splitter

The light path in a camera indicator system is bent
by means of Dichroic-reflectors as shown in Figure 4.
Several effects occur in this type of installation;
namely,ghost images,polarization of the reflected light
and modification of the light over the image-plane.Al-
though all may be of interest, here,the discussion will
center on the polarization and band-splitting results.
For a CRT with P-7 or P-14 phosphor,ray 1 will contain
both a blue flash and a long-persistence yellow light,
each with equal components of horizontal and vertical
polarizations.In Figure 4 the horizontal components are
normal to the paper while the vertical components are
in the plane of the paper and perpendicular to the
light rays.
At the dichroic coating the incident light divides in a
manner determined by the nature of the coating and the
inclination of the beam-splitter.Analysis shows that

the reflected light contains more horizontal than vertical polarization components for reflection favors light components in the plane of the reflecting surface.Tracing through the full optical system is beyond this paper.What is important here is that polarization filters,grids,and anti-reflection are used in the optical design.They have also find wide application in antenna designs.

7.0 Design Similarities

Geometric optics provide a common thread linking antenna and optical design.

Geometric optics has been applied,with success, in designing and analyzing reflector-feeds,reflector-imaging problems,and lens design.

Frequency independent antennas offer a classical example where designing from nearly dc to daylight is possible.In this type of antenna all of the electrical dimensions are independent of frequency.If no restriction in size exists,a very low frequency design can be provided.If an optically-fine feed-point is made designs approaching 40 to 50 Ghz may be realized.

As a final point,again,consider the loop and Hall sensor.Both have used ferrite to enhance performance.

8.0 Concluding Remarks

Geometric optics or physical optics have found use in both optics and antennas.

Anti-reflective coatings,absorptive materials,grids and filters have all found use in optics and antennas.

References

[1]G.J. Monser, " Pickup Devices for Very-Low- Frequency Reception",Electronics,pp-,1961.

[2] D.Levine and G.J.Monser,"Polarization,Ghost,and Shading Effects in Dichroic Beam-Splitters",J.Opt.Soc. Am.,51:783-789,1961.

[3] H.Gent," The Bootlace Aerial",Ministry ofTechnology,England,1958.

[4]W.Rotman and R.F.Turner,"Wide-angle Microwave Lens for Line Source Applications",AFCRL-68-18,1962.   W.and Turner,R.F. (1962) Wide-angle Microwave

 Lens for Line Source Applications,AFCRL-68-18,1962. AD666053

Table 1 Field Intensity to Give Unity Output

Operating Frequency (CPS)	Bandwidth (CPS)	Input in Microvolts/Meter		
		Hall Sensor	Loop	Whip
10,000	10	2400	0.15	0.13
1,000	10	2400	1.50	1.32
20	10	5200	76	66
10	10	5300	153	132

Notes: 1.Hall sensor,InSb with flux collectors.
2.Loop,H=5inches,W=15 inches,2 turns.
3.Whip,H=15 inches.

V, Hall output terminals

B

1"

1"

Ferramic Q

Bias Current, I

InSb wafer
3/8 x 1/8 x 0.01 inches

Ferramic Q
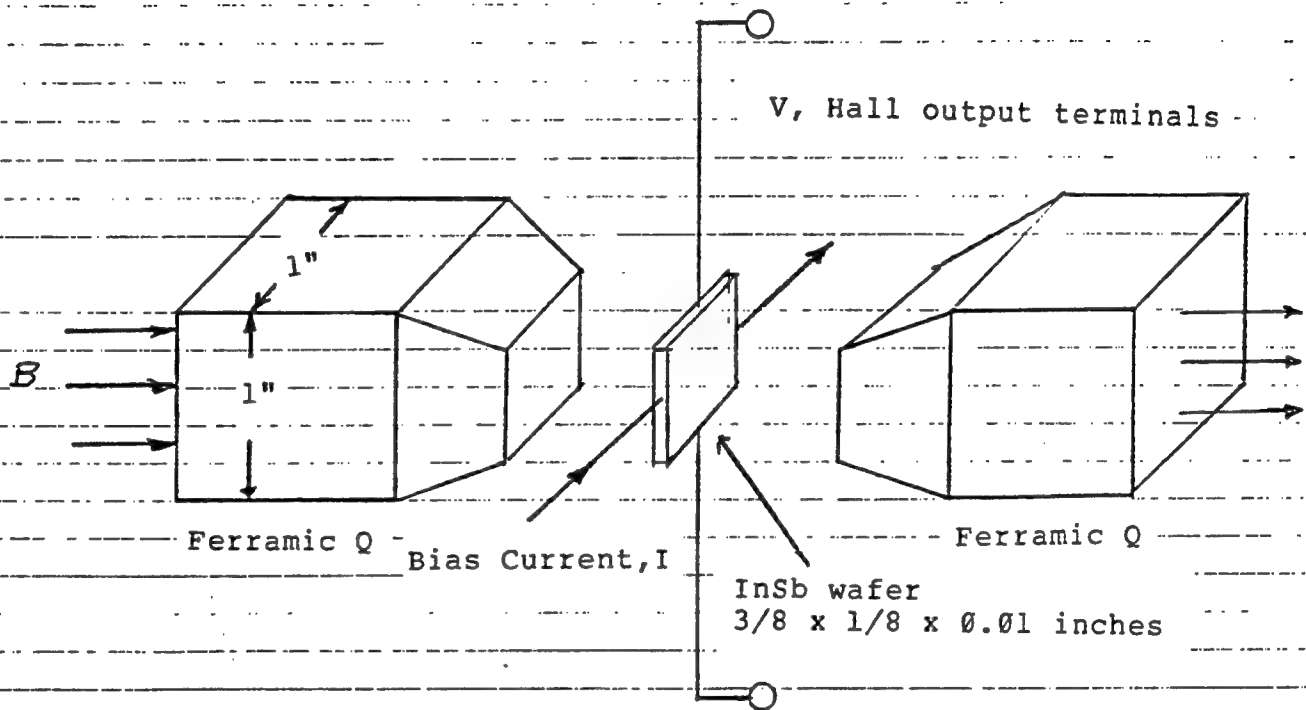
Notes: 1. In use the ferrite faces touch the sensor.
2. Face dimensions of ferrite match the wafer
dimensions.

Figure 1 Hall sensor with Flux Collectors
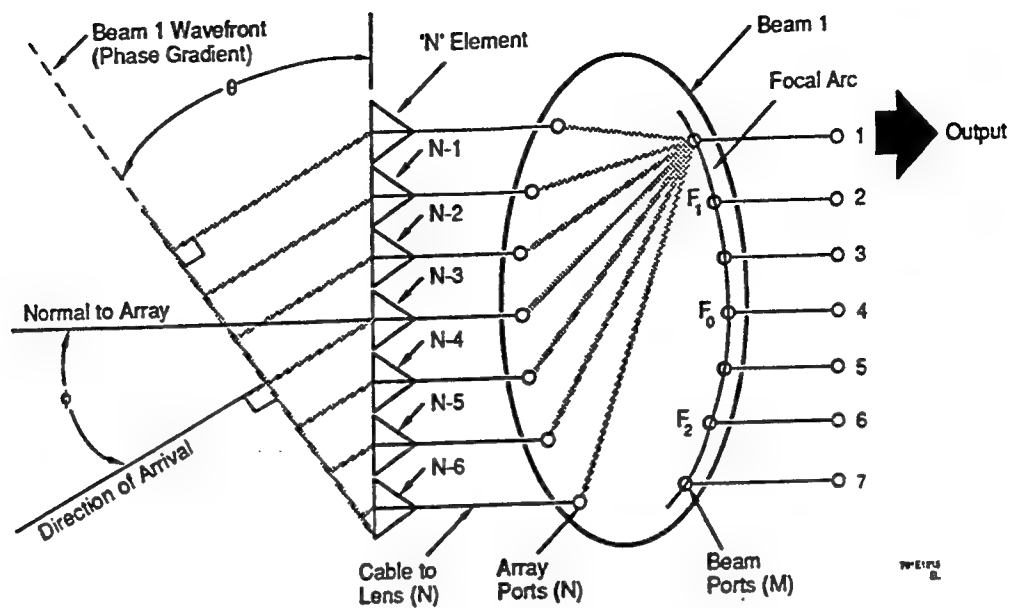
Figure 2 Rotman Lens

Grid wires on thin dielectric sheet
spacing,0.1 wavelengths or less

S_2
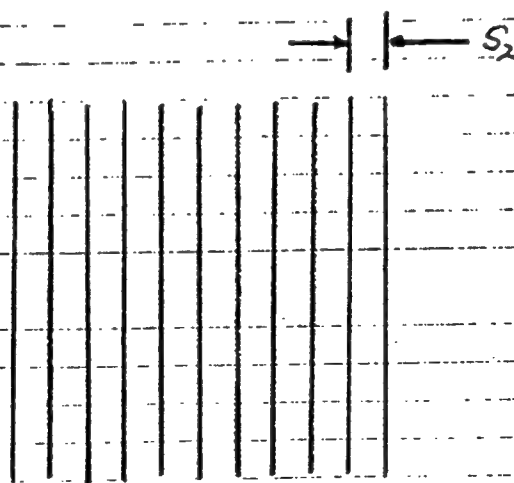
Figure 3 Grid Structure
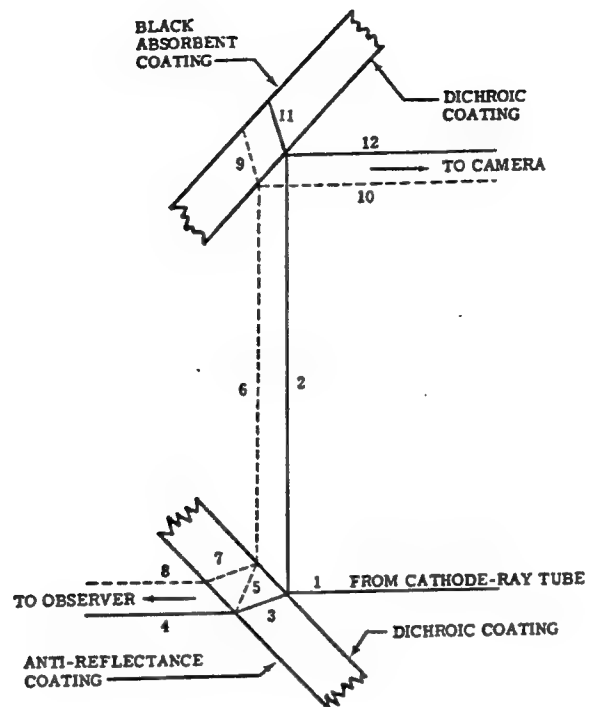
Figure 4 Dichroic Beam-splitter,
Typical Radar Application
(Courtesy J.Opt Soc.Am.)

THE ELECTRONICALLY TUNED HELIX[*]

George Ploussios
GS Inc.
60 Central Street
Andover, MA 01810

ABSTRACT: Normal mode helices that are electronically tuned (ETH) are described. These electrically short (as small as .029λ) monopole designs are electronically tuned by shorting turns of the helix, making the helix resonant in the 1/4 wavelength mode at the operating frequency. At this frequency the input impedance is nominally 50 ohms (no base tuner). Since the inherent RF losses of the ETH are very low, it is a much more efficient antenna (in the order of 10 db better at the lowest operating frequency) than designs of equal size.

The theory of operation is presented. Design equations and curves are used to determine the antenna parameters. The AWAS method of moments computer analysis program is then used for fine adjustments of these design parameters. The design equations are presented and tradeoffs between antenna efficiency, Q, and size are discussed.

Operation over frequency bandwidths in excess of 15:1 has been realized. Test results are presented. The test units are 1) a 12 inch tall ETH in a blade configuration that operates over the 30-88 MHz VHF band, and 2) a 10.5 foot ETH with 9 foot top loading whip that operates over the entire 2-30 MHz HF band. The test results are in agreement with theory.

1.0 INTRODUCTION

The Electronically Tuned Helix (ETH) is an efficient electrically small element [1] tunable over a decade or greater frequency bandwidth. It is particularly suited for applications in the HF through UHF bands as a monopole or dipole.

The schematic of the ETH monopole, a quarter wave resonant helix, is illustrated in Figure 1. The helix is shorted to ground at the base and is fed at a fixed point above ground. Tuning is accomplished by changing the radiating

_____

[*]Patent # 4,924,238

structure to make it resonant at the operating frequency. This is done by electrically shorting helix turns with PIN diode switches. When there are no turns shorted the helix resonates at its lowest frequency. Increasing the number of turns shorted increases the element resonant frequency. The radiation pattern from the ETH monopole corresponds to that obtained from conventional monopoles of similar size.

The ETH element impedance is real at resonance, varying from zero at the antenna base to a value approaching infinity at the antenna top. A feed point along the helix is selected that nominally corresponds to the 50 ohm point throughout the tunable operating band.

The diode switch tuning arrangement has several desirable features. These include:

1) The bias wires for the PIN diode switches are embedded in the helix windings. The DC return path is the helical radiator which is shorted to ground at the antenna base. Shielding the bias lines from the antenna fields eliminates any field induced RF currents in the bias network.

2) The RF voltage on the helix winding at the bias wire exit point and the RF voltage at the PIN diode bias feed point (points A and B in Figure 1b) is the same. This results in no RF coupling (zero RF current) to the bias feed lines.

3) The balanced, cathode to cathode or anode to anode PIN diode switch configuration is consistent with good RF switch design practice that minimizes signal distortion. The result is very low spurious harmonic content.

The described bias network design obviates the need for series feed line RF chokes and therefore: a) avoids feed line RF losses; b) avoids the generation of non radiating circuit resonances; and c) avoids instabilities caused by variation in the choke impedance at antenna resonance.

The ETH antenna's small size, high efficiency, high Q, and wide tuning bandwidth make it an appealing design choice for systems that do not require a

wide instantaneous bandwidth. These characteristics are summarized below.

Size: Practical ETH designs are typically 3 to 15 times smaller than conventional resonant antennas at the lowest operating frequency. Examples include the 13.5" 30-90 MHz ETH SINCGARS antenna tested by the Army on an AH-64 APACHE helicopter and a 2-30 MHZ Army shelter mounted 10.5' ETH with top loading whip.

Antenna Q: The Q of the ETH is inherently high. Qs well is excess of 100 (measurement at 3 dB points) at the lowest operating frequency have been experimentally measured. High Q antennas may require impedance tracking circuitry and tuning compensation in order to avoid instability due to changes in temperature, installation dimensions, humidity, and the physical environment. In order to avoid this circuitry, a maximum Q between 50 and 100 is typically specified for the ETH. In some designs lowering the antenna Q to a maximum of 50 requires the intentional introduction of RF loss. As frequency increases from the minimum, Q will decrease without any help from the designer.

Efficiency: Since the ETH element always operates at resonance, its radiation resistance is significantly higher than that obtainable from an equivalent sized electrically short monopole, i.e. 60% higher [2]. Also, since the antenna does not require any additional matching components (e.g. base loading coils or matching network), the only losses in the system are those associated with the helix surface loss and diode losses, both of which are low. The net result is that ETH antennas will exhibit little or no loss at frequencies at which the antenna is greater than $.07\lambda$ long in the case of a monopole, or $.14\lambda$ long in the case of a dipole. At lower frequencies, Q limiting losses designed into the system will reduce antenna efficiency. In addition, as the electrical length decreases, the antenna radiation resistance rapidly decreases, making the winding and diode losses more significant. However, theoretical and test data indicate that the smaller than $.07\lambda$ ETH will still exhibit an efficiency that is at least 10 dB higher than that obtained from conventional designs of comparable size and operating bandwidth.

Tuning (operating) Bandwidth: The ETH tuning is accomplished through the use of PIN diode switches that exhibit low loss and high switching speed.   Tuning over many frequency octaves has been demonstrated, e.g. HF and VHF designs tested by the Army cover 15:1 (2-30 MHZ) and 3:1 (30-88 MHZ) frequency bands. Less than 10 μsec. switching speeds have been measured with the ETH.

2.0 DESIGN PARAMETERS

The initial antenna design parameters are selected using the design equations and curves developed over the past 8 years by the author. The AWAS method of moments analysis program is then used to 1) improve on the design accuracy and 2) help determine the tuning table that relates operating frequency to switch state.

The helix is a slow wave structure. Its axial phase velocity relative to the free space phase velocity determines the antenna size reduction factor. Specifically for the basic unshorted helix mounted on an infinite ground plane:

$$s = c/v = \sqrt{1 + 20 \, (N \, D)^{2.5} \left(\frac{D}{\lambda}\right)^{.5}} \qquad (1)$$

where: s = antenna size reduction factor
c = velocity of light
v = phase velocity along the helix axis
D = helix diameter
N = helix winding turns per unit

The helix length is then computed to be $\dfrac{\lambda_o}{4 \, s}$ where λ_o is the wavelength at the lowest operating frequency, f_o. The other pertinent equations at this frequency are:

$$E_o = \frac{40}{40 + s^2 R_{loss}} \qquad (2)$$

$$Q = \frac{s^3 E_o K}{102} \qquad (3)$$

where E_o = Efficiency at the lowest frequency

R_{loss} = series antenna system loss resistance

$K = 60 [\ln \frac{4H}{D} - 1]$

$Q = 3$ dB Q at f_o

The helix winding resistance is the principal source of loss in the ETH. Losses in the PIN diodes are secondary, assuming low loss diodes are used and are properly biased. The loss resistance attributable to the above sources is typically a few tenths of an ohm. This plus any intentionally added resistance make up the value R_{loss}.

From equation 3, Q increases rapidly as s increases. At about s = 4 the designer needs to consider the introduction of network loss as a means to limit Q to the 50 to 100 specified maximum. This reduces the efficiency of the antenna at f_o. However, as frequency increases efficiency rapidly improves. The nominal expression for efficiency vs. frequency is:

$$E = \frac{E_o}{(1 - E_o)(f_o/f)^p + E_o} \qquad (4)$$

where the value p is dependent on the tuning curve selected and details of the design. Usually p is greater than 2.5.

460

Several ETH HF and VHF designs have been built and tested. Two of these designs are discussed below.

3.0 ETH DESIGNS

3.1 30 - 88 MHZ VHF ETH

The VHF ETH, shown in figure 2, consists of a tunable helix with an elliptical cross section designed to fit in a blade shaped radome. The total height of the antenna package is 13.5" which includes a 12" tall helix (.03λ at 30 MHz). Eight PIN diode tuning switches, capable of creating 2^8 tuning conditions, are mounted along the surface of the helix.

The VHF ETH was designed for installation on the tail or belly of the APACHE helicopter and operates with the ARC-201 SINCGARS radio. It has been tested at GS Inc. and by the Army at Lakehurst and Ft. Monmouth NJ.

VSWR: The VSWR of this design, installed on an 8 foot diameter ground plane, is illustrated in Figure 3. The curve shown consists of overlapping VSWR plots obtained for selected (total of 82) tuning states. This unit is capable of being tuned over a 27 MHz to 100 MHz frequency band. The maximum Q over the 30 to 88 MHz band occurs at 30 MHz and was measured to be 50. Relatively lossy helix winding material used in this design limited the maximum Q achieved.

The same antenna was tested mounted on top of the APACHE tail (Figure 4). The distance between the antenna and tail rotor was only 12". This made antenna tuning very difficult at frequencies near the rotor resonance, which occurred in the 45 to 50 MHz region. The VSWR obtained in this installation is illustrated in Figure 5 (26 tuning states). The Q at 30 MHz was measured to be 13, much lower than in the flat ground plane installation. This is attributable to an increase in radiation resistance caused by the currents generated on the tail surface.

Gain: Initially, relative gain measurements[*] were made comparing the gain of a breadboard version of the above described ETH to the 54" AV-405 and the 84" AV-412 antennas mounted on an 8 foot ground plane. The results obtained are illustrated in Figure 6. The gain measurements were made with the test antenna at ground level and the receive terminal on top of a building that was 1/4 mile away.

The calculated gain of the bread board ETH tested is illustrated in figure 7 for an infinite ground plane installation. The gain, based upon published data, of the AV 405 and the AV 207-2, a 14.5 inch tall antenna with a trailing top load, is also illustrated in that figure. The measured relative gain shown in Figure 6 is in reasonable agreement with the postulated gain curves in Figure 7.

Airborne measurements using the ETH shown in Figure 2 and the 84 inch AV-412, both mounted at the same location on the APACHE tail , were taken at Lakehurst and at Ft. Monmouth. Gain measurements were taken at Ft. Monmouth at a test site that is used for measuring relative antenna gain. The Lakehurst tests were run to demonstrate ETH SINCGARS operation, to obtain pattern data, range data and at least qualitative gain data.   Both test sites produced measurement data that indicated substantially greater ETH gain at the lower frequencies.

The gain measurements taken at Ft. Monmouth were taken with the APACHE airborne and at four different azimuth orientations. At each orientation, frequency was swept and signal level recorded.   The difference between the signal levels obtained using the ETH and AV-412 were averaged for the four orientations and plotted as relative gain in Figure 8.   The improvement in relative gain shown for the ETH in the Figure is consistent with the tests discussed above and illustrated in Figure 6.

Measurements taken at Lakehurst with the ETH and AV-412 were taken on different days, under different weather conditions and without a common reference. Therefore accurate gain measurements were not possible. However, measurements taken at Lakehurst on these antennas did indicate substantial gain improvement at the lower frequencies when using the ETH, and comparable gain

---

[*]The airborne measurement program was directed Alexander Strugatsky, ATCOM, US ARMY, St. Louis, MO.

at the higher frequencies . Sample patterns taken are illustrated in Figures 9 and 10. At the higher end of the VHF band the installation geometry had a greater impact on the ETH pattern than on the AV-412 pattern. This is understandable since the AV-4112 is a dipole that is nominally decoupled electromagnetically from the tail, whereas the ETH tested is a monopole dependent on the tail as its ground plane.

<u>Power:</u> The antenna was designed to handle 20 watts CW. It has been tested with the SINCGARS radio and with a 20 watt CW source.

3.2 HF ETH

The HF ETH, shown in Figure 11, is a 10.5 foot tall 6.5 inch diameter helix with 15 PIN diode tuning switches. Without the 9 foot top loading whip the antenna is tunable from 2.7 MHZ to greater than 30 MHz (antenna height of $.029\lambda$ at 2.7 MHz). With the top loading whip in place the lowest tuning frequency is 2 MHz (height of $.04\lambda$).

The ETH antenna was designed to be collocated with a second ETH of the same design on an 8' by 12' by 8' shelter. The antenna is hinged at the base and is capable of being lowered onto the shelter top and the top loading whip removed. The lowering to the stowed position as well as erection can be accomplished by one person.

<u>VSWR:</u> The measured VSWR of this antenna mounted on a 8' by 12' by 8' metal building is shown in Figure 12. A total of 97 tuning states were used to obtain the tuning curve illustrated. The Q at 2 MHz was 20 and is lower than the computed Q assuming an infinite ground plane. The lower Q may be attributable to the extra height the shelter presents.

Above 13 MHz the antenna becomes electrically long, i.e. $> 1/4 \lambda$ high. At this point, one can choose to operate using higher order modes, e.g. 3/4 mode, or attempt to continue in the 1/4 mode by attenuating the current along the upper portions of the structure. The latter approach was pursued for the tests used in obtaining the data for figure 12. Attenuation of the upper portions of the ETH is accomplished by shorting the helix from the base up to a point where an LC trap

decouples the upper portion of the antenna. This allows operation in the 1/4 wave mode from 2 MHz to > 30 MHz with the broad elevation pattern associated with the 1/4 mode operation.

Gain: Calculated antenna gain for an infinite ground plane installation is shown in Figure 13. Measured gain data on the shelter is in the process of being reduced and will be available at a future date.

Power: The HF ETH was designed for operation at the 40 watt level. The power limitation on the ETH antennas is a function of the diodes selected and the bias voltages used. Power levels in the KW region are feasible using available components.

4.0 CONCLUSIONS

The ETH antenna concept featuring very small efficient antennas tunable over bandwidths exceeding 15:1 has been demonstrated in the field. Specifically a 13.5 inch tall airborne VHF ETH has been developed and tested. The gain of this antenna has been measured to be higher than antennas 4 to 6 1/2 times larger; 6 to 14 dB higher in the 30 to 40 MHz band. The gain relative to antennas of comparable size is estimated to be > 10 dB higher throughout the VHF band. In addition, an efficient HF antenna consisting of a 10.5 foot tall ETH plus a nine foot top loading whip covering the 2 to 30 MHZ range has been built and tested. Two of these antennas separated by 10 feet were installed on a shelter roof.

5.0 ACKNOWLEDGEMENTS

6.0 REFERENCES

[1] G. Ploussios, "An Electronically Tuned Helical Antenna", *Microwave Journal*, May 1991

[2] A. G. Kandoian and W. Sichak, "Wide Frequency Range Tuned Helical Antennas and Circuits", *IRE Convention Record 1953 part 2*

$$\frac{V_1 + V_2}{2}$$

(b) ETH Diode Switch Detail

(a) ETH Monopole

HELIX

PIN DIODES

BIAS FEED LINE

RF FEED
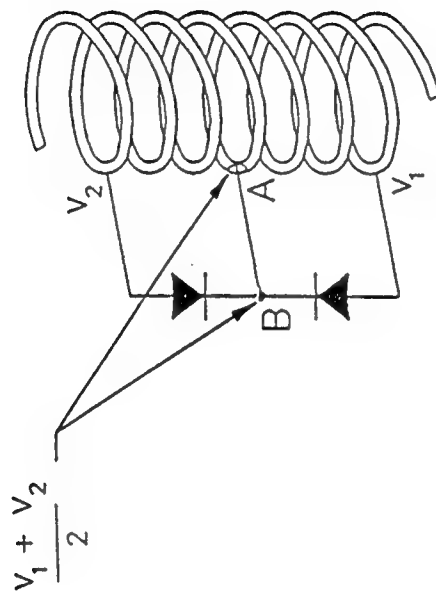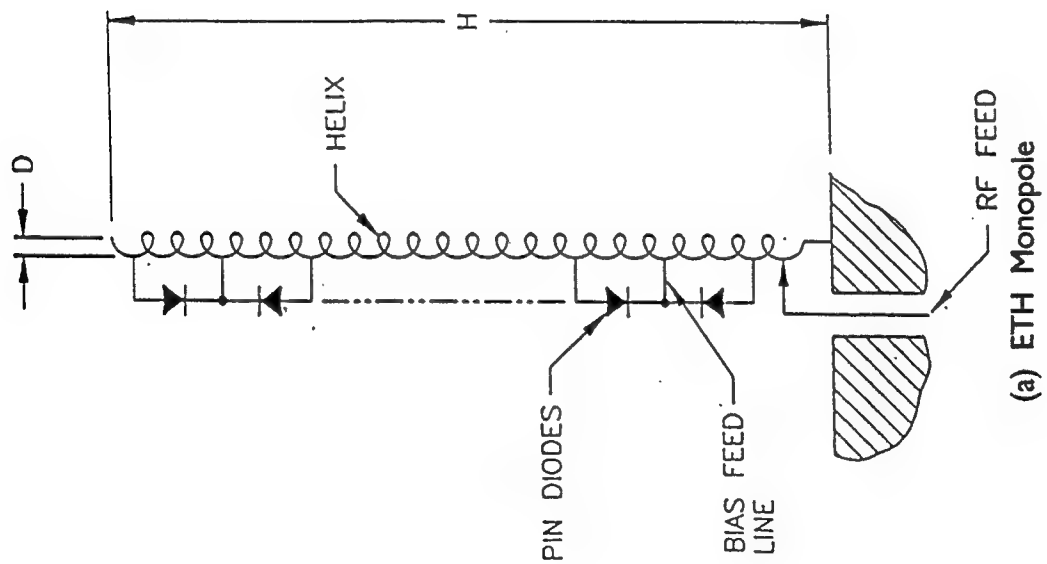
Figure 1 ETH Schematic

Figure 2 VHF ETH

467
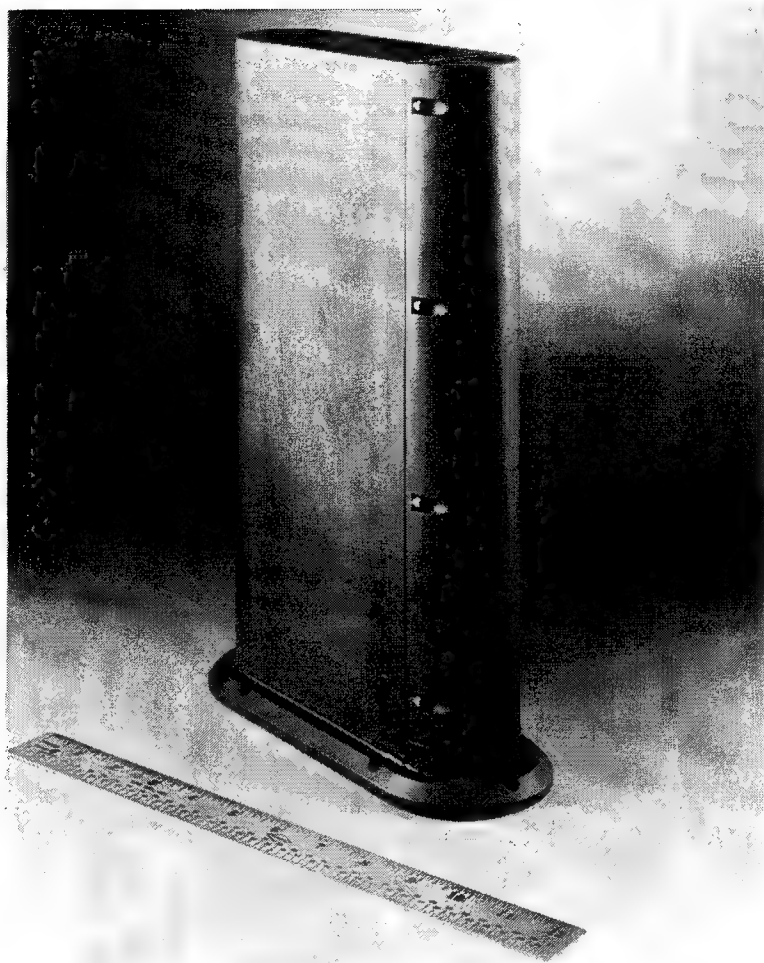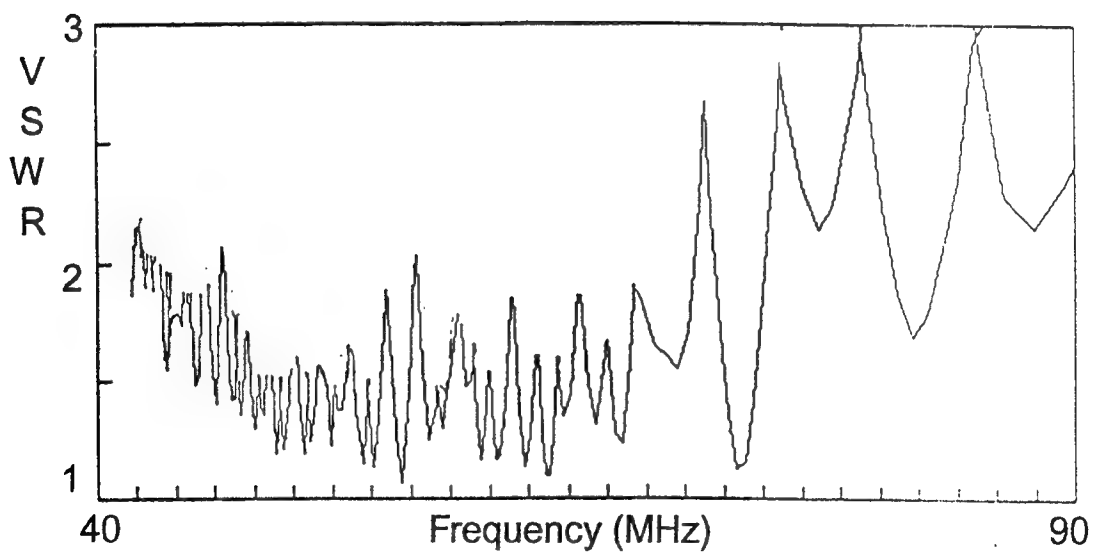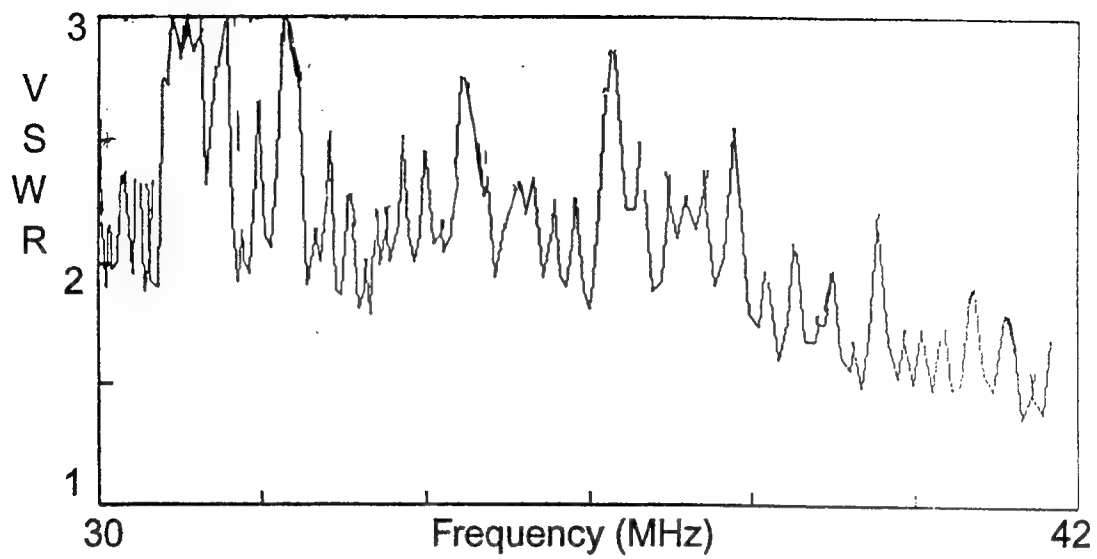
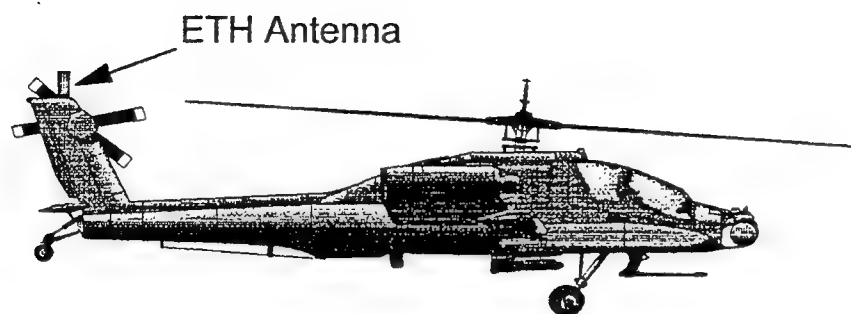Figure 3 VSWR of VHF ETH Installed on 8 Foot Ground Plane

Figure 4 ETH AH-64 APACHE Installation
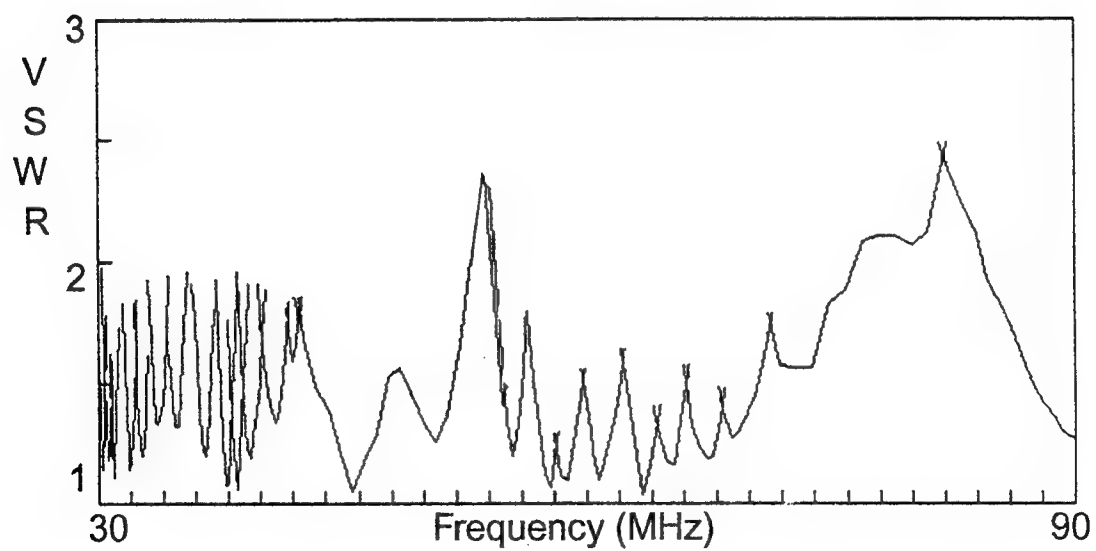
Figure 5 VSWR of VHF ETH Installed on APACHE Tail
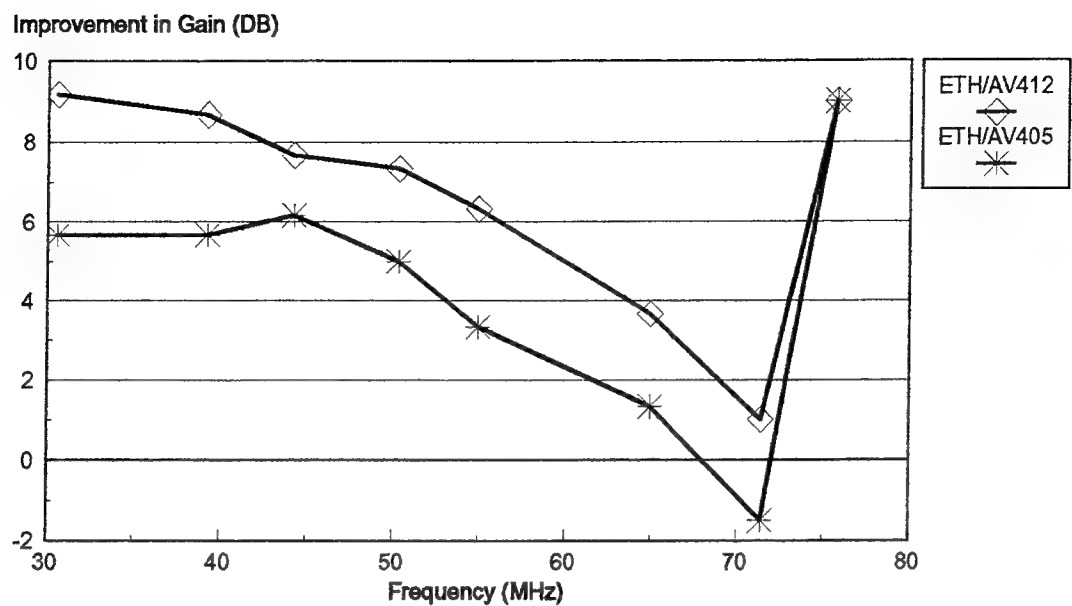
Figure 6 Measured Breadboard ETH Gain Relative to AV-412 and AV-405
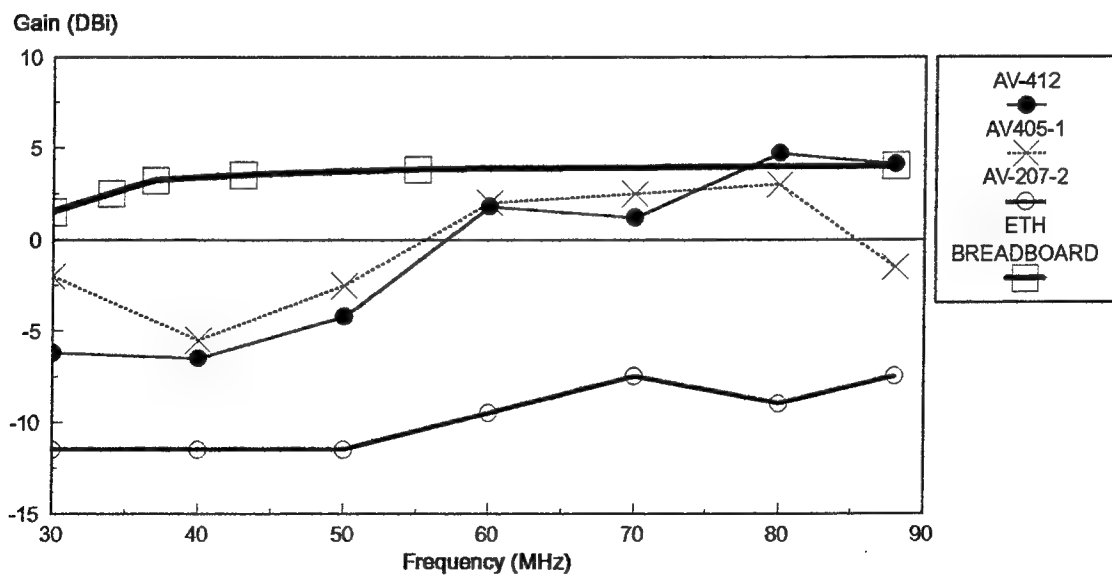
Figure 7 Gain of ETH, AV-412, AV-207-2 on Infinite Ground Plane and AV-412

470
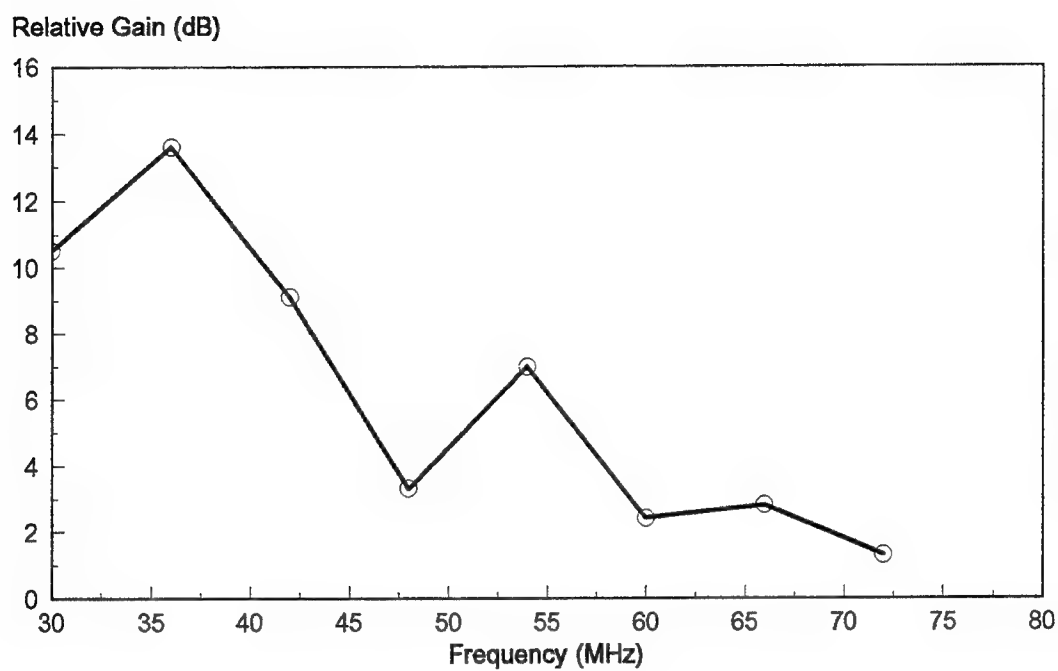
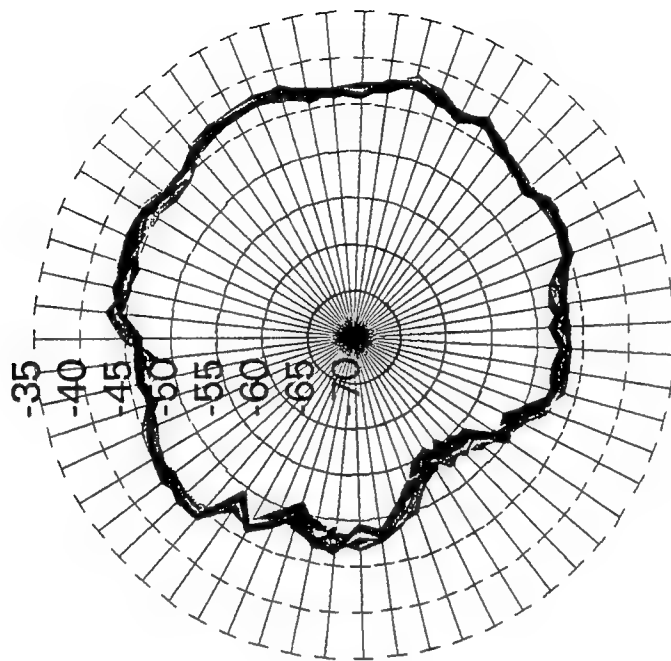Figure 8 Measured ETH Gain Relative to AV-412 - APACHE Tail Installation
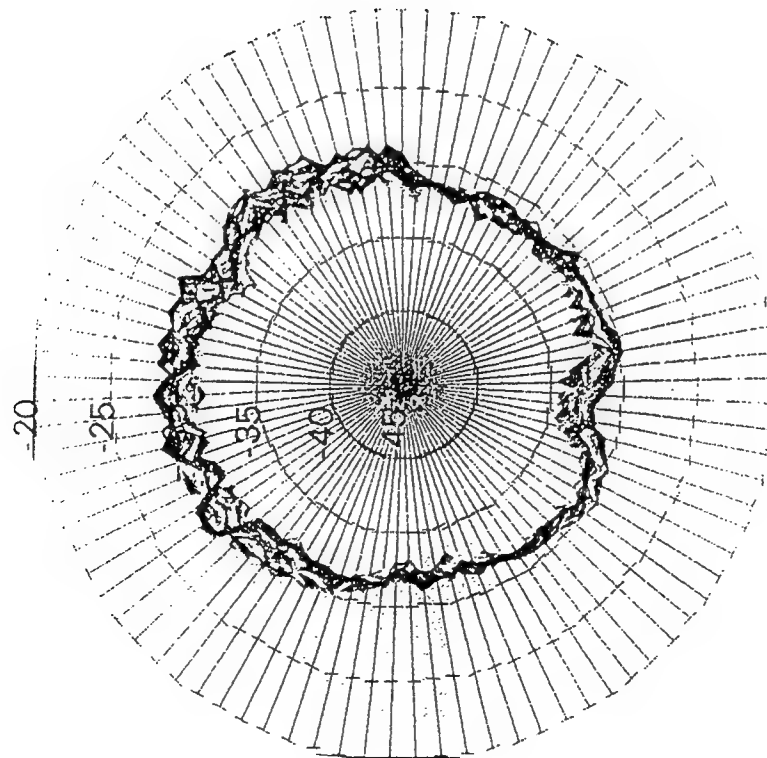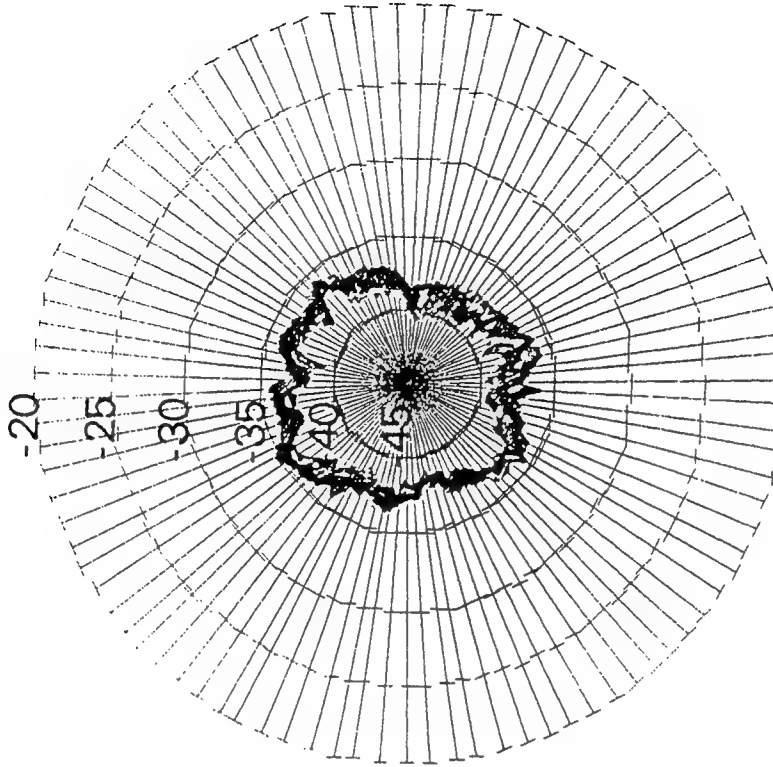
b) AV-412

a) ETH

Figure 9 Pattern Measurements at 33.725 MHz
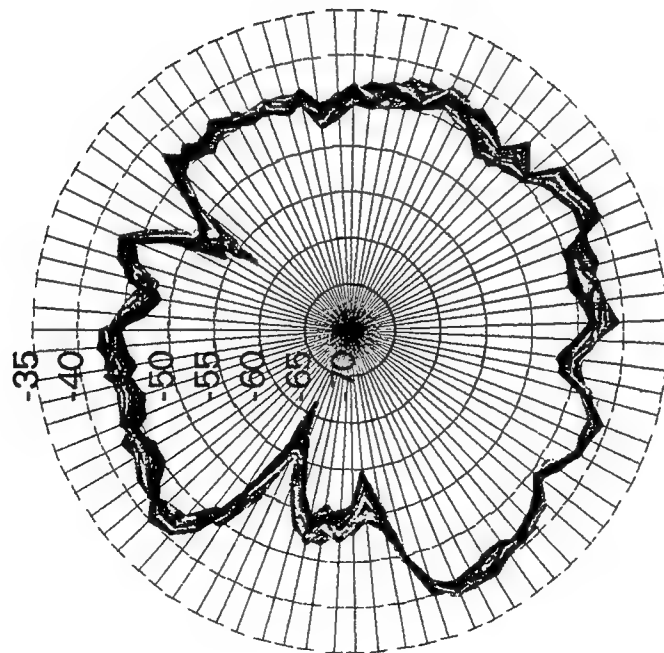
a) ETH

b) AV-412

Figure 10 Pattern Measurements at 71.4 MHz
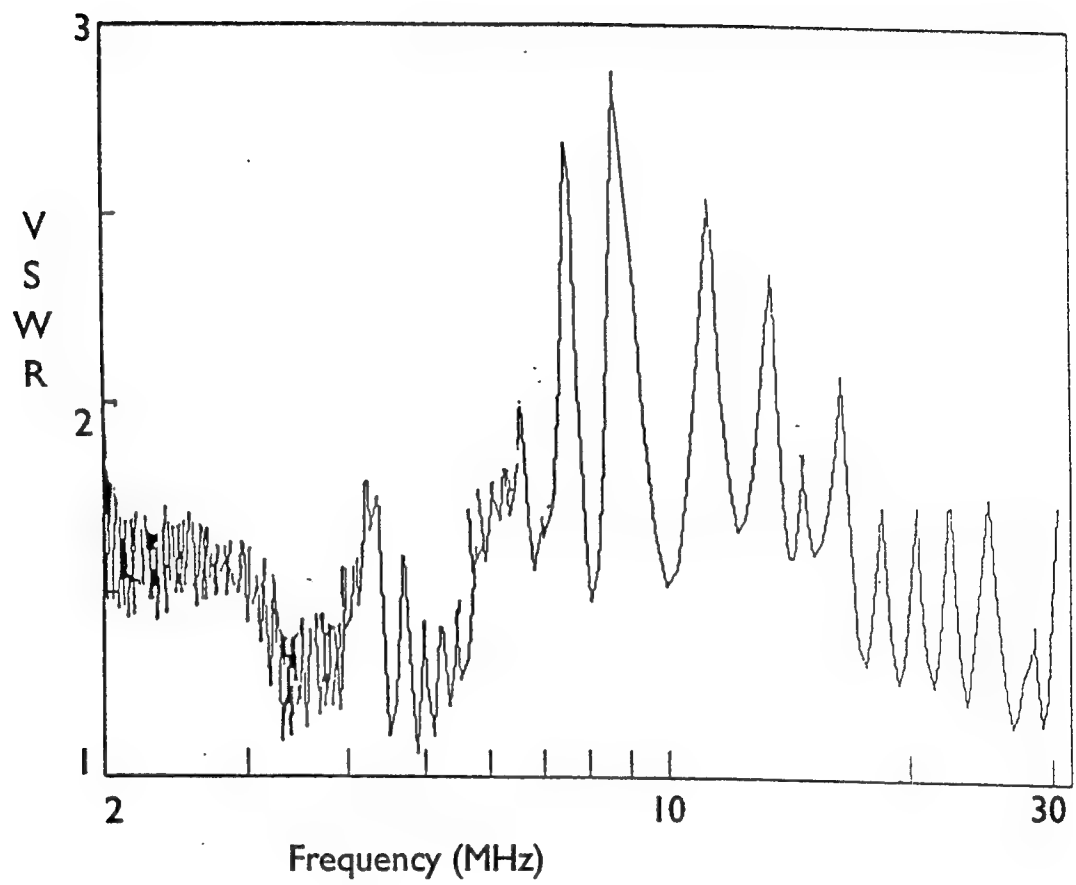
Figure 11 HF ETH

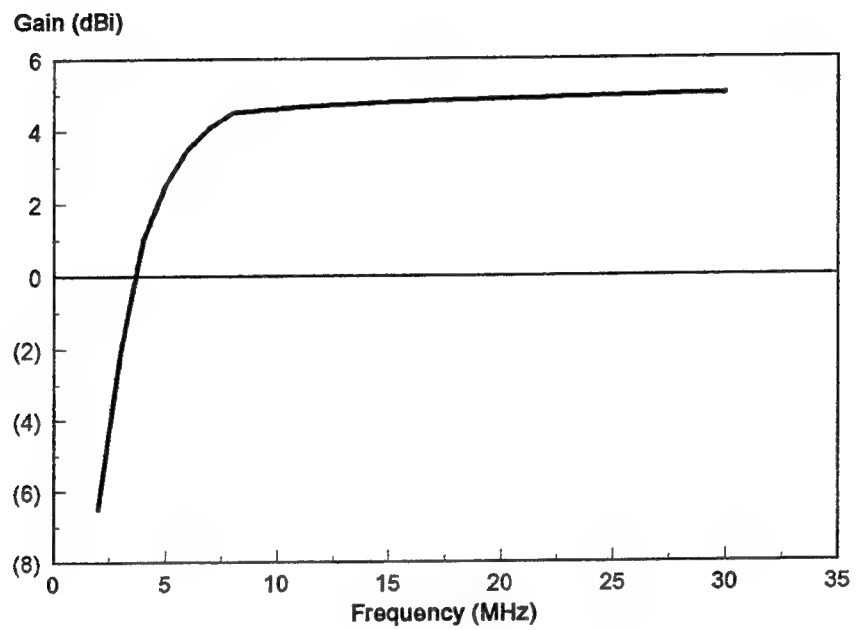Figure 12 HF ETH VSWR vs. Frequency

475

Figure 13   HF ETH Gain -   Infinite Ground Plane Installation

HEMISPHERICAL COVERAGE USING

A LOADED MONOPOLE

EDWARD E. ALTSHULER
U.S. AIR FORCE
ROME LABORATORY
ELECTROMAGNETICS AND RELIABILITY DIRECTORATE
HANSCOM AFB, MA. 01731-3010

ABSTRACT

A traveling - wave distribution of current can be produced on a monopole over a ground plane by inserting a resistance of suitable magnitude one-quarter wave length from its end. It has been shown that the resistor can be replaced with a modified folded dipole which has a radiation resistance approximately equal to that of the resistor. Furthermore, it was found that when the folded element was .1λ above the ground plane, near hemispherical coverage was obtained in the plane of the element. In this paper the resistor is replaced with orthogonal folded elements. The main objective is to obtain an antenna which is circularly polarized and which has near hemispherical coverage. Input impedance, current distribution, and radiation patterns of this antenna are computed using the Numerical Electromagnetics Code (NEC); radiation patterns are also measured.

In order to produce a circularly polarized wave it is necessary to tilt the folded elements. After experimenting with a number of different configurations an antenna design having directional and polarization properties approaching those that were desired was obtained. All computations and measurements were conducted at 1.6GHz, a frequency close to that of the proposed IRIDIUM Communications System and the L1 Band of Global Positioning System (GPS). The Double-Folded Monopole may prove to be very low cost option as a ground antenna for these systems.

1. INTRODUCTION

A traveling-wave distribution of current can be produced on a linear antenna by inserting a

resistance of approximately 240Ω one-quarter wavelength from its end [1]. The main

disadvantage of the resistive-loaded traveling-wave antenna is that it is only about 50% efficient

because part of the input power is absorbed by the resistor. It is possible to replace the resistor

with a resonant antenna having a radiation resistance approximately equal the matching resistor [2]. Thus the input section still has a traveling-wave distribution of current up to the inserted element, but now the power previously dissipated in the resistor is also radiated. In a previous paper [2], the antenna replacing the resistor was a modified folded dipole. It was shown that when the inserted folded element was approximately .1 wavelength above the ground plane, near hemispherical coverage could be obtained in the plane of the folded element. In this paper, the resistor is replaced with orthogonal folded elements in order to achieve both circular polarization and near hemispherical coverage. The input impedance, current distribution and radiation patterns of this loaded monopole over an infinite ground plane are computed using the Numerical Electromagnetics Code (NEC) [3]. Radiation patterns are also measured, but over a finite ground plane. This limitation leads to a discrepancy between computed and measured patterns. This antenna can be designed to operate at any frequency as long as it does not become physically too small. All calculations and measurements were made at 1.6GHz, a frequency close to that of the proposed IRIDIUM Communications System and the L1 band of the Global Positioning System (GPS). The double-folded monopole [4] has the potential of being a low cost ground antenna for these systems.

2. APPROACH

The procedure that was used in this investigation was to first calculate the antenna properties using NEC, construct the antenna, measure its characteristics and then compare the computed and measured results where possible. The NEC computations were only used as a guide for the antenna design; obtaining good agreement between computations and measurements was of

secondary concern. The NEC is a computer code for calculating the electromagnetic response of antennas. Modeling a wire structure involves both geometrical and electrical factors. Each wire is divided into segments that are generally within the range from 0.01λ to 0.1λ; shorter and longer segments should be avoided. The circumference of the wire should be very small compared with the wavelength and the wires should not be too close together. It was found that the input impedance was very dependent on the segment lengths of each wire; thus a sensitivity analysis was conducted. The input impedance was computed for a wide range of segment lengths. The segment length around which the input impedance showed the least variability was considered optimum. The radiation pattern was not significantly affected by the segment length. The input impedance and radiation patterns were most sensitive to the overall lengths of the modified folded dipoles; the other lengths, wire diameter and tilt angle were not as critical.

The NEC output is a tabulation of the following data.

- The input impedance and admittance along with the power radiated for a one volt source,

- The real and imaginary components along with the magnitude and phase of the current for each segment of the antenna

- The far-field radiation pattern including the vertical, horizontal and total gain for each elevation and azimuth angle; also the magnitude and phase of the E_θ and E_ϕ fields, the sense of polarization, axial ratio and tilt angle of the major axis.

The proposed antenna has three mutually perpendicular radiating elements, the two horizontal crossed modified folded dipoles and the vertical monopole. The crossed dipoles provide coverage

in the zenith direction and higher elevation angles; the monopole coverage is primarily at the low elevation angles. Ideally we would like the currents in the orthogonal elements to be in phase quadrature so that a circularly polarized field is radiated; this is obviously not possible for all directions. We have chosen to concentrate on obtaining right-hand elliptical polarization down to elevation angles as low as possible. A general expression for the power received by an elliptically polarized antenna from an elliptically polarized transmitting antenna is

$$\frac{P_R}{P_0} = \frac{1}{2} \pm \frac{2r_1 r_2}{(1+r_1^2)(1+r_2^2)} + \frac{(1-r_1^2)(1-r_2^2)}{2(1+r_1^2)(1+r_2^2)} \cos 2\alpha \qquad (1)$$

where

r_1 is the axial ratio of the receiving antenna

r_2 is the axial ratio of the transmitting antenna

α is the angle between the major axes of the two ellipses of polarization

P_r is the power received

P_0 is the maximum power

The (+) sign is used if the senses of polarization are the same while the (-) sign is used if the senses are opposite. If we assume that the transmitting antenna is right-hand circular polarization then $r_2 = 1$, $\alpha = 0$ and (1) becomes

$$P = \left[\frac{1}{2} \pm \frac{r_1}{1+r_1^2} \right] \qquad (2)$$

The polarization loss in dB is $\quad 10 \log_{10} \left[\frac{1}{2} \pm \frac{r_1}{1+r_1^2} \right] \qquad (3)$

and is plotted in Fig.1. Note that as long as the receiving antenna has the same sense polarization as the transmitter then the maximum polarization loss of 3dB occurs when the receiver is linearly polarized. If however, the receiving antenna has the opposite sense polarization, then the polarization loss becomes very large. Thus, it is of utmost importance to limit the presence of opposite sense polarization in the receiver.

Many configurations of the double-folded monopole were investigated. Interesting results were obtained but both right-and-left-hand elliptical polarizations were generally present if the folded elements were confined to a horizontal plane. If however the folded elements were tilted with respect to the horizontal plane a more dominant single sense of polarization could be produced for most directions. A design that produced good results is similar to that shown in Fig.2. The first folded element fed from the monopole was in the x-plane and tilted downward. The upper segment of that fold then fed the x- and y- plane folds,in parallel and were titled upward. The y- plane fold fed the other y- plane fold which was tilted downward and then both x- and y- plane folds intersected on the z- axis and fed the end section of the monopole. The modified folded dipoles were tilted (up or down) at an angle of about 40° with the horizontal; they had lengths of .44λ. The lower and upper parts of the monopole each had a height of .25λ; thus the total half-length of the monopole was .50λ;. The wire circumference was approximately .016λ.

The NEC computations were done for a set of radiation patterns for the θ and ϕ planes for both E_θ and E_ϕ polarizations over an infinite ground plane. These data were generally sufficient to assess the performance of the antenna. The antenna was constructed from #18 copper BUS wire

481

which has a diameter of about 1.0mm. However, in order for the wire to hold its desired shape it had to be stretched; this process reduced the wire diameter slightly. The antennas were hand made using a jig.

Although the antenna dimensions were only approximately those that were used for the NEC computations they were believed to be close enough to make a qualitative comparison. The antenna was mounted over a 1.2 m x 1.2 m (6.5λ x 6.5λ)ground plane. All data were collected at a frequency of 1.6Ghz. Directional properties were measured in an indoor pattern range.

3. RESULTS

The directive gains were computed for the double-folded monopole for the E_θ and E_ϕ components for both the θ- plane and ϕ- plane. For illustration in Fig.3 we show the gain for the E_θ and E_ϕ components as a function of zenith angle at an azimuth angle of $\phi = 0°$. We also show the total gain $E = \sqrt{E_\theta^2 + E_\phi^2}$ for the same angles. Finally we show the gain of this antenna for a right-hand circularly polarized incident wave. Note that the total gain and circularly polarized gain are reasonably flat except at the low elevation angles. Antenna patterns at other azimuth angles in the θ plane were similar to those in Fig. 3. In Fig. 4. we show the computed gains for the E_θ and E_ϕ components in the ϕ - plane for a zenith angle of $\theta = 60°$. As in Fig. 3 the corresponding total gain and circularly polarized gain are also shown. It is seen that the total gain is very uniform and the circularly polarized gain with the exception of slight depressions at azimuth angles near 90° and 300° is also quite uniform in the azimuth plane. A complete set of computed circularly polarized patterns in the θ - plane are shown in Fig. 5 for azimuth angles of 0°, 45°, 90° and 135°. A corresponding set of patterns in the ϕ - plane for zenith angles of 50°,

60° and 70° is shown in Fig. 6. It is seen that circularly polarized coverage is good down to elevation angles of about 30°; when the elevation angle is further decreased to 20° we see a null develop near an azimuth angle of ϕ =90°. Corresponding circularly polarized patterns were measured in the θ - plane for azimuth angles of 0°, 22.5°, 45°,67.5°, 90°,112.5°, 135° and 157.5°. For clarity only the patterns for ϕ=0°, 45°, 90° and 135° are shown in Fig. 7. It is seen that the pattern shifts to the right with increasing azimuth angle; the patterns for the intermediate angles which were omitted from Fig. 7 fall in their expected locations. Unfortunately, it was not possible to measure the patterns in the θ - plane since the pattern range was not set up for this measurement. However, the patterns in Fig.7 do confirm that the circularly polarized coverage is good down to elevation angles near 30° from the horizon as predicted from the computations. In Figs. 8-11 the computed and measured circularly polarized patterns are compared. It is seen that the comparison is excellent. It can be shown that the ripples in the measured pattern and the loss of signal near the horizon arise from the fact that the measurements were made using a finite ground plane whereas the computations were done for an infinite ground plane. The computed input impedances were found to be somewhat sensitive to the segment incremental lengths. In Fig. 12 the input impedance is plotted as a function of the incremental length which was normalized to wavelength. It is seen that the impedance is least variable for a length of approximately .022λ which produced a value of 72.4 +j 7.8 ohms.

4. CONCLUSION

The objective of this investigation was to show that it is possible to obtain circularly polarized coverage over most of the hemisphere using a double-folded monopole. The initial results are

very encouraging. The agreement between computed and measured patterns was very good considering that infinite and finite ground planes were used spectively. Also, the fabricated double-folded monopoles were rather crude and the dimensions were not exact. There are still many variations of the double-folded monopole that have not been tested; it is possible that better coverage can be obtained. Future plans are to examine the frequency dependence of the antenna. Also, a horizontal loop will be used in place of the modified folded elements to obtain the $E\phi$ component. Whereas many circularly polarized antennas that are used for hemsipherical coverage require complex phasing networks, the loaded monopole is fed directly from a coaxial line and may provide a very low cost option for achieving the desired coverage.

5. ACKNOWLEDGEMENT

6. REFERENCES

[1] E. E. Altshuler, "the traveling-wave linear antenna," IRE Trans. Antennas Propagat., vol. 9, pp. 324-329, July 1961.

[2] E. E. Altshuler, "A monopole antenna loaded with a modified folded dipole," IEEE Trans. Antennas Propagat. vol., 41, pp.871-876, July 1993.

[3] G. J. Burke and A. J. Poggio, "Numerical Electromagnetics Code (NEC) - Method of Moments," Rep. UCID18834, Lawrence Livermore Lab. CA, Jan. 1981.

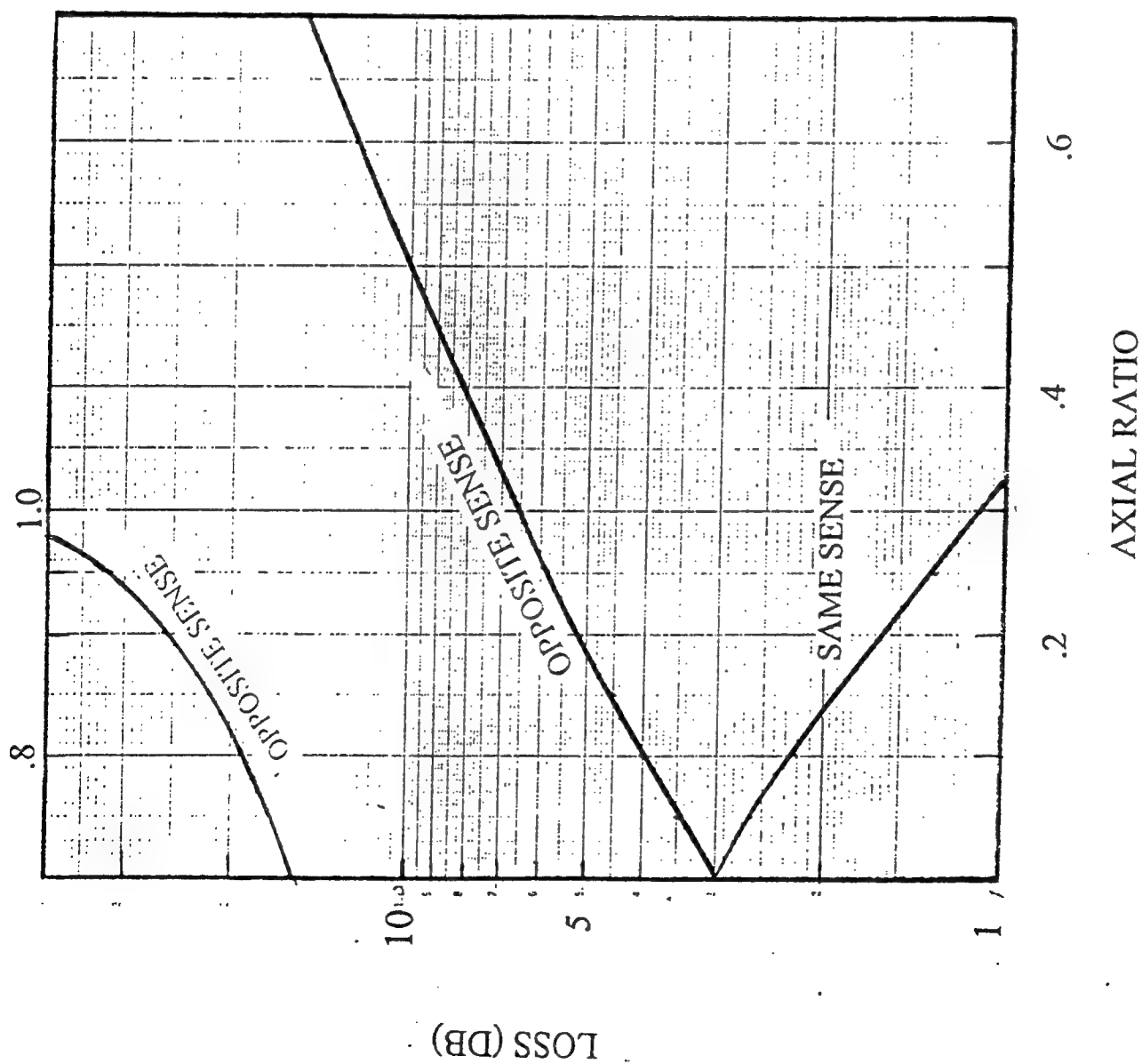[4] E. E. Altshuler, Double - Folded Monopole, U.S. Patent #5,289,198, Feb. 22, 1994.

FIGURE 1. POLARIZATION LOSS
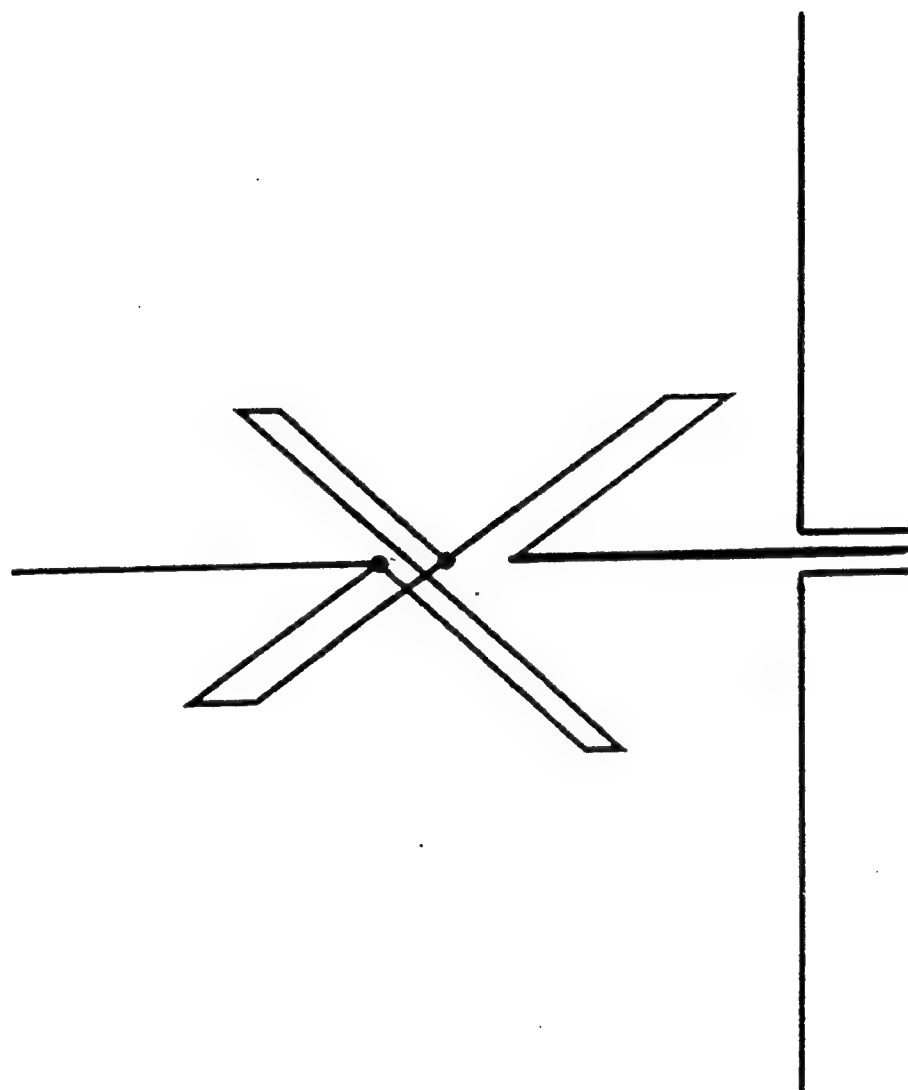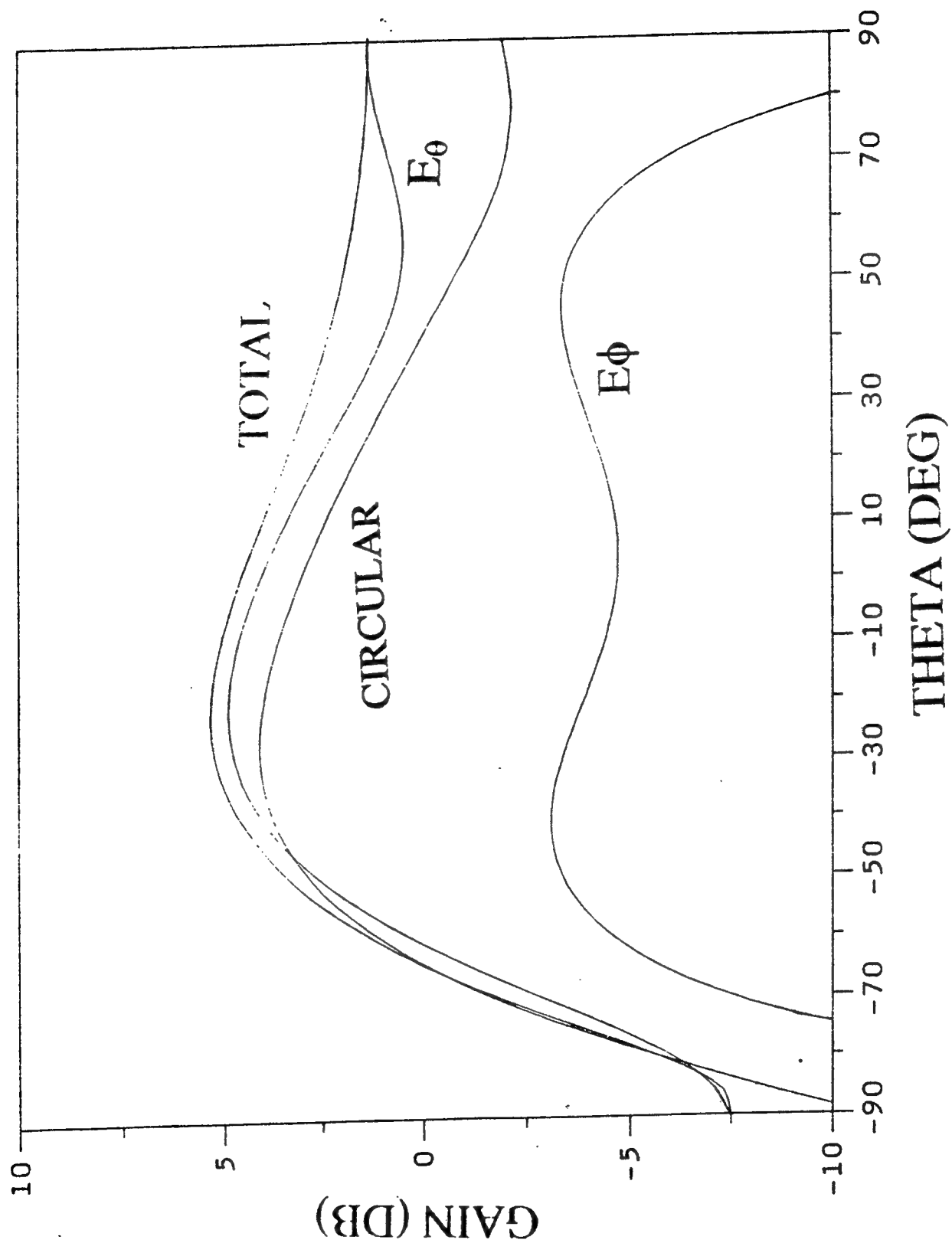
FIGURE 2. DOUBLE-FOLDED MONOPOLE
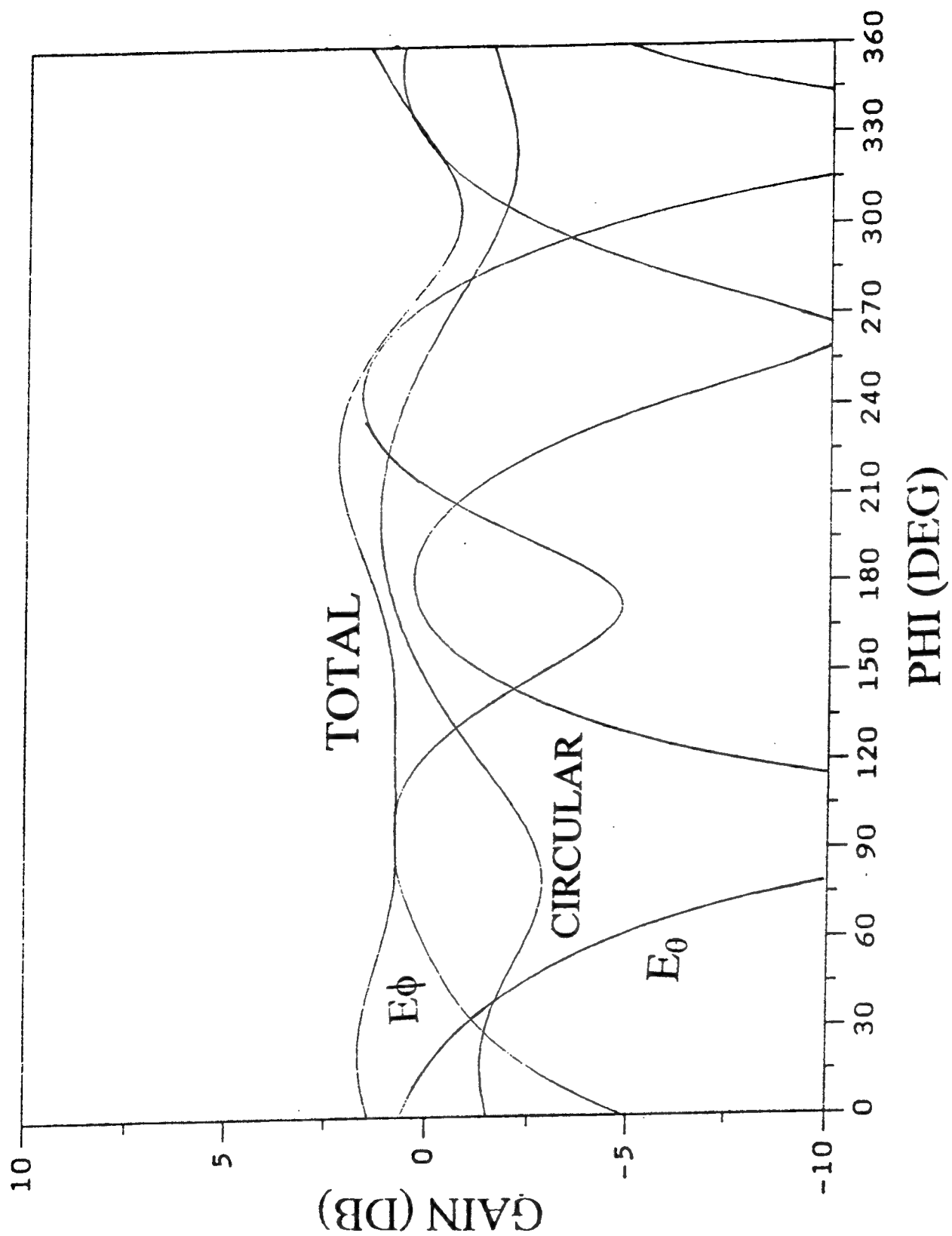
FIGURE 3. COMPUTED PATTERNS IN θ - PLANE (φ = 0°)

FIGURE 4. COMPUTED PATTERNS IN φ - PLANE ($\theta = 60°$)
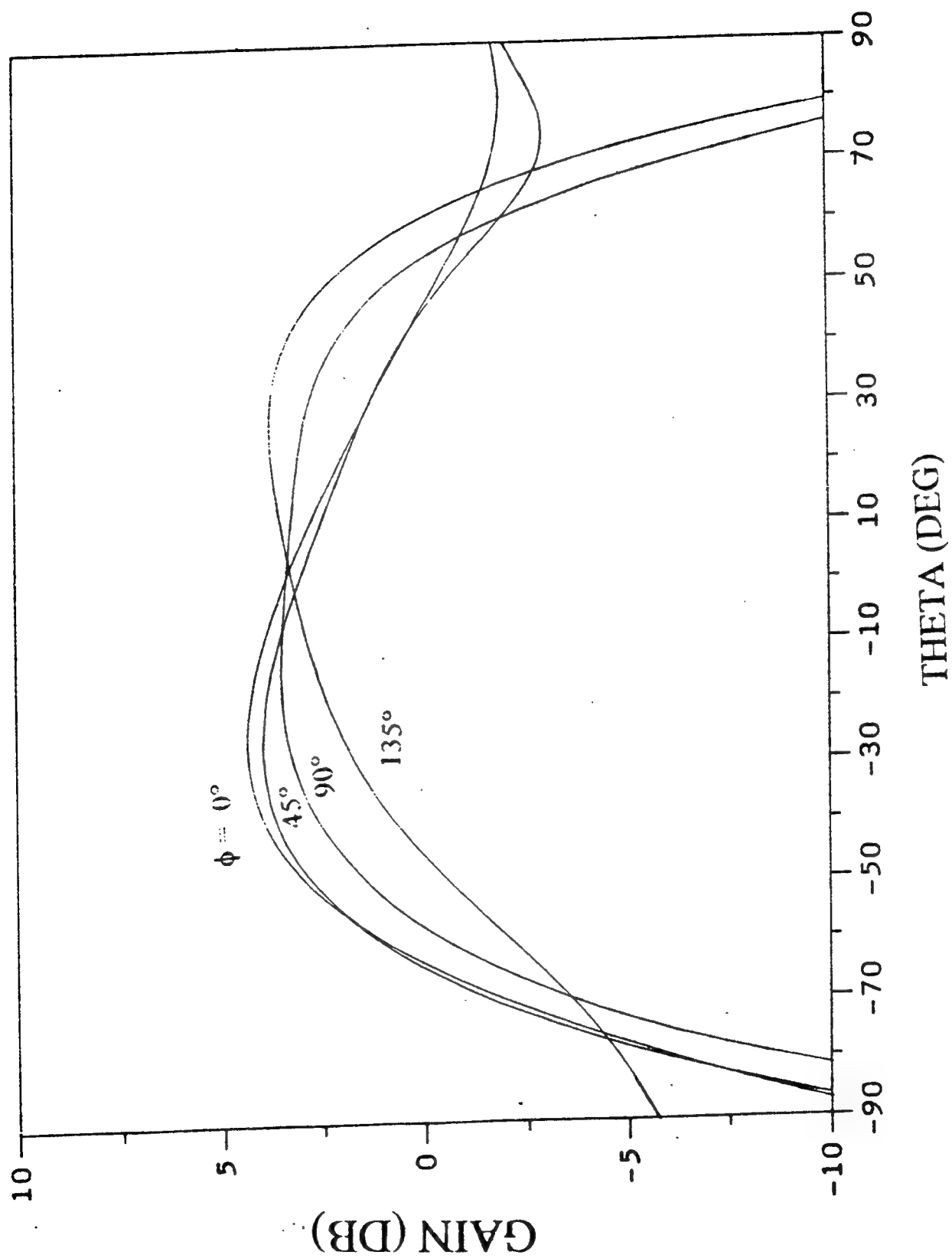
488

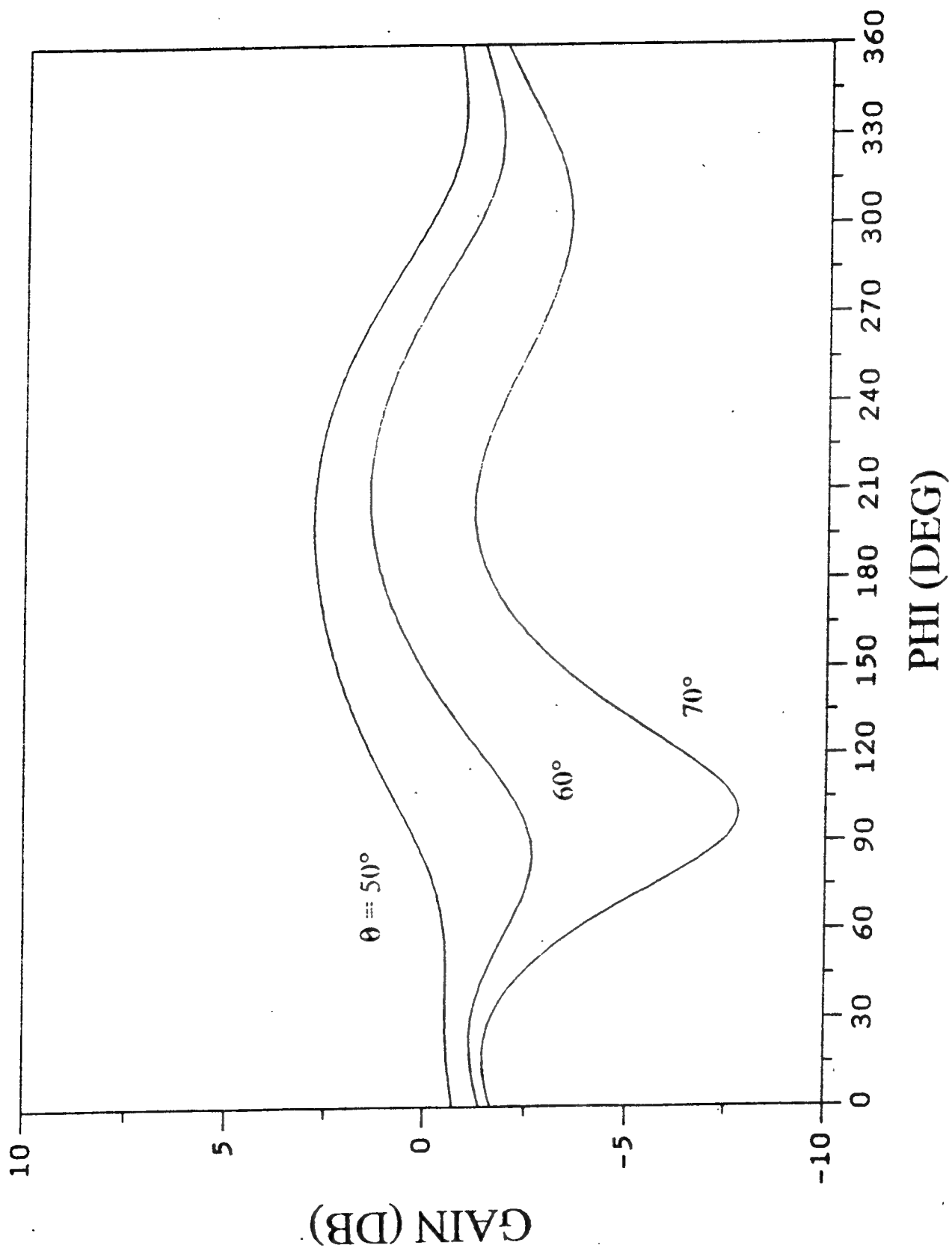FIGURE 5. COMPUTED CIRCULARLY POLARIZED PATTERNS IN θ- PLANE
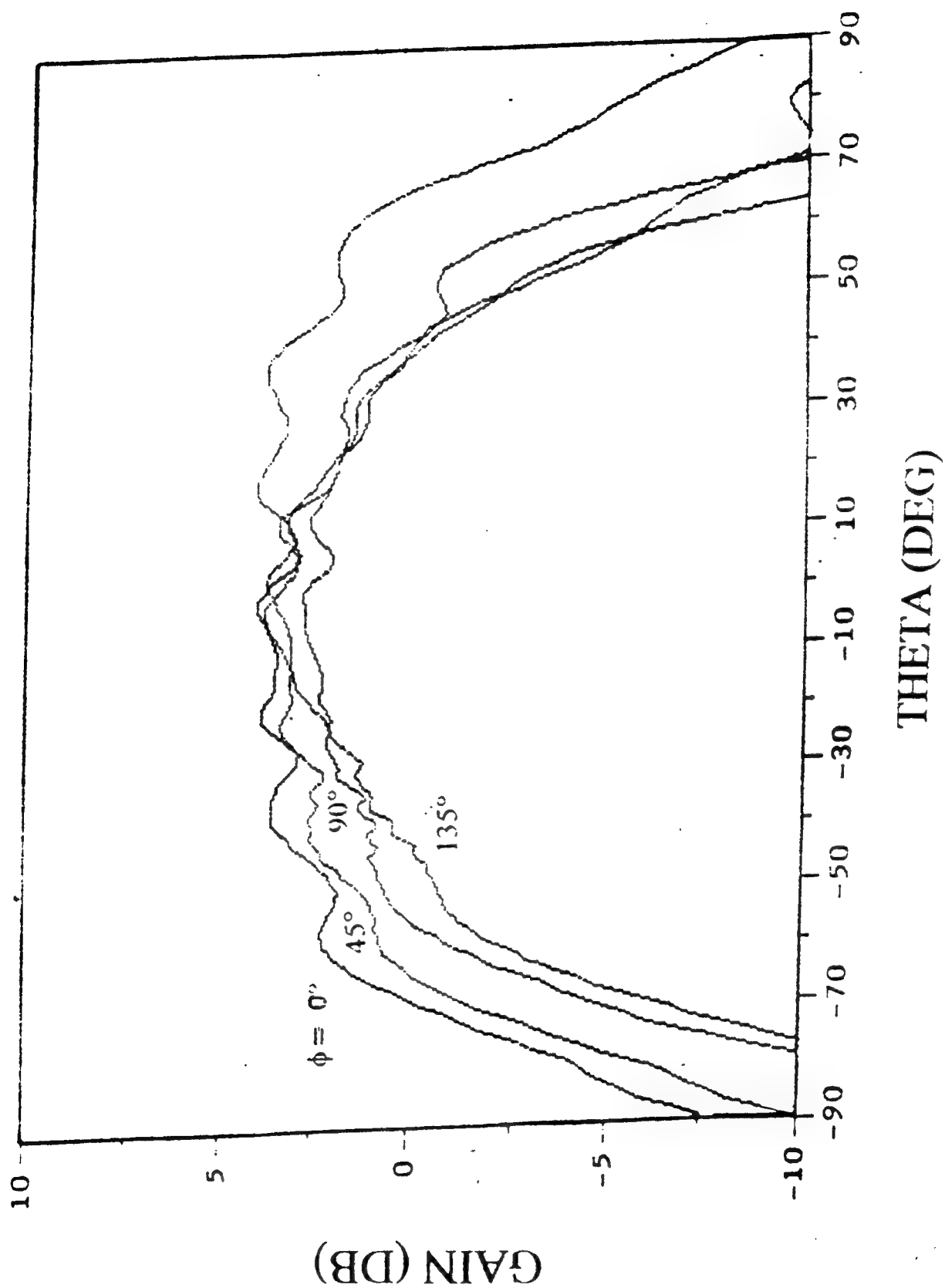
FIGURE 6. COMPUTED CIRCULARLY POLARIZED PATTERNS IN φ - PLANE

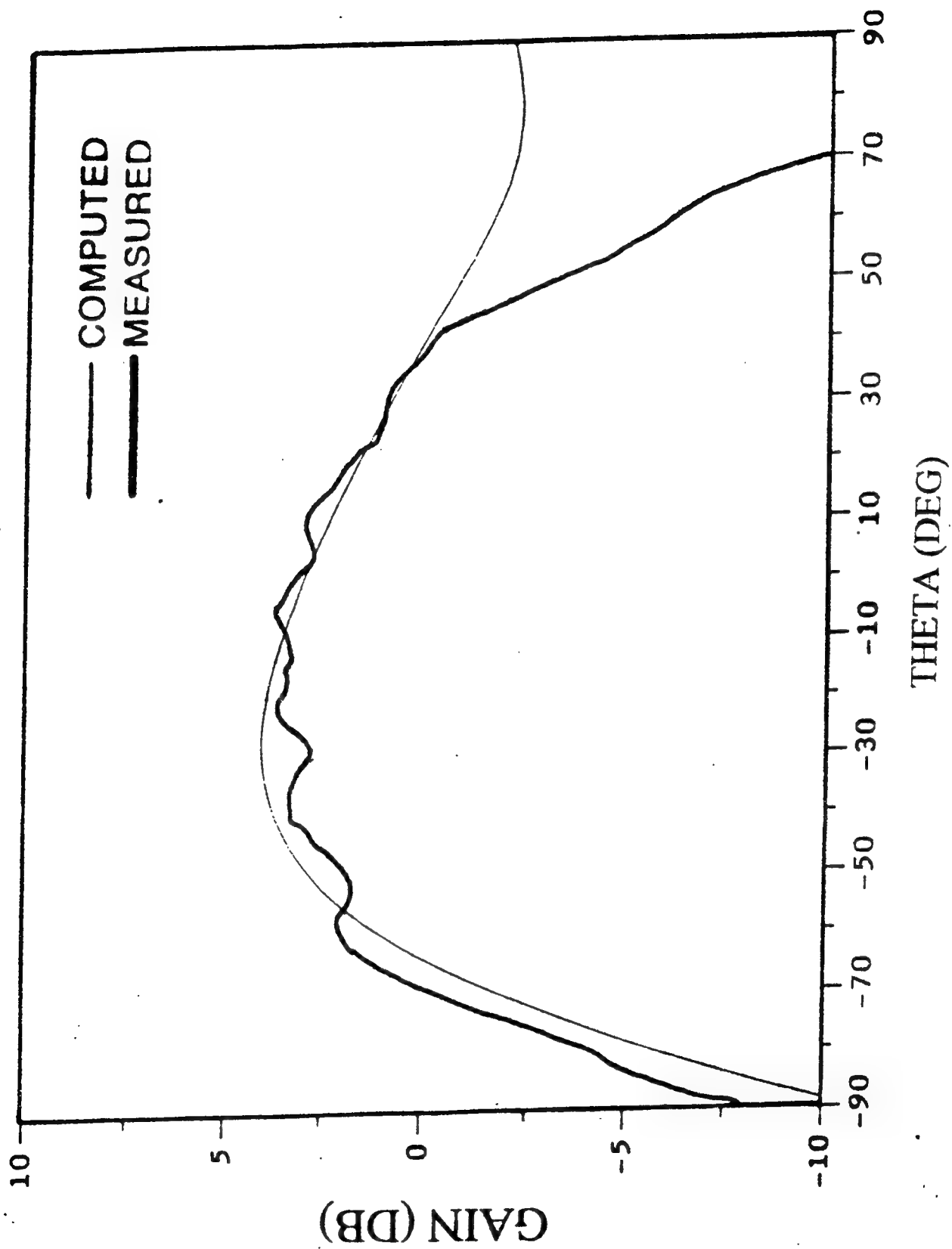FIGURE 7. MEASURED CIRCULARLY POLARIZED PATTERNS IN θ - PLANE

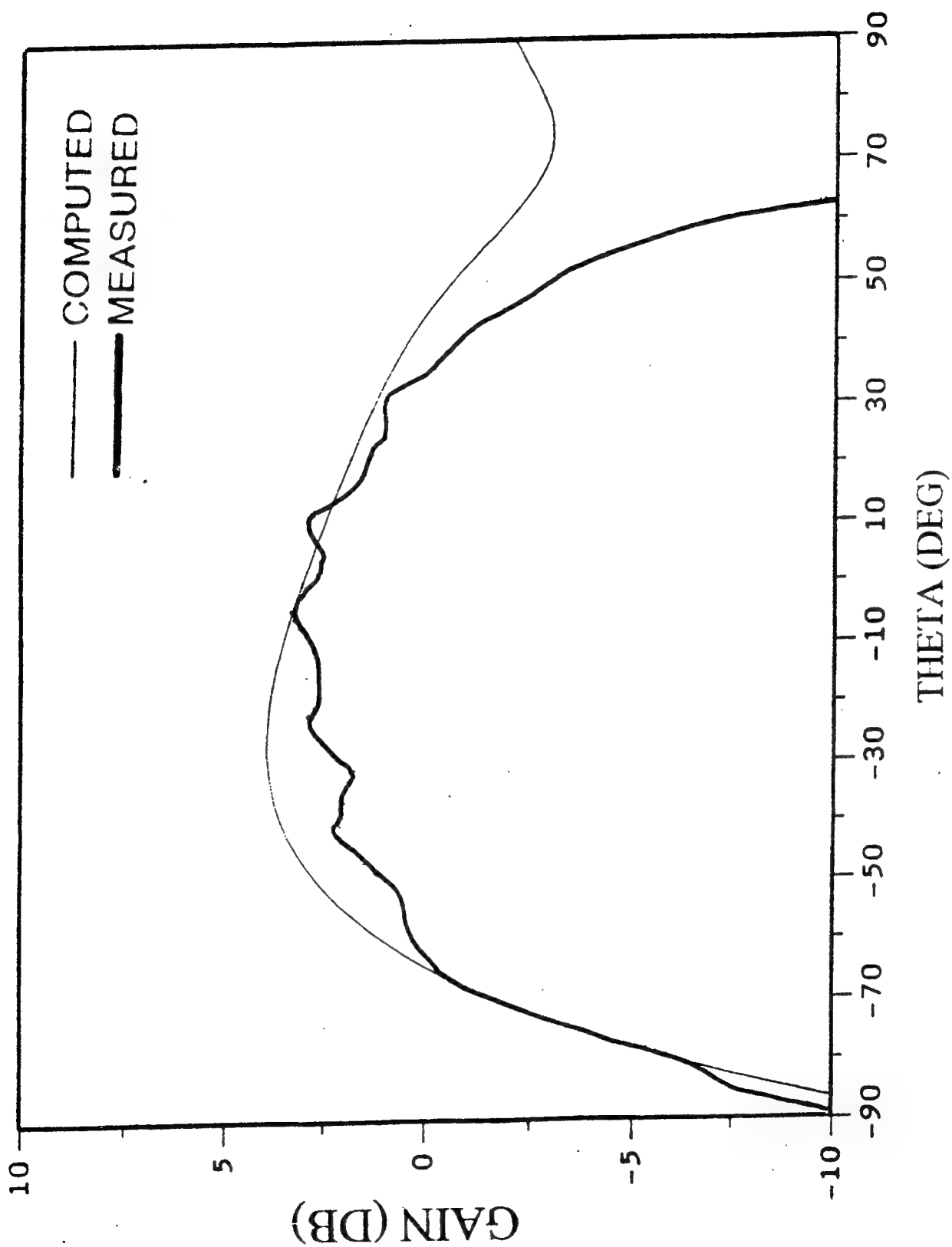FIGURE 8. CIRCULARLY POLARIZED PATTERNS IN θ - PLANE (φ = 0°)

FIGURE 9. CIRCULARLY POLARIZED PATTERNS IN θ - PLANE (φ = 45°)
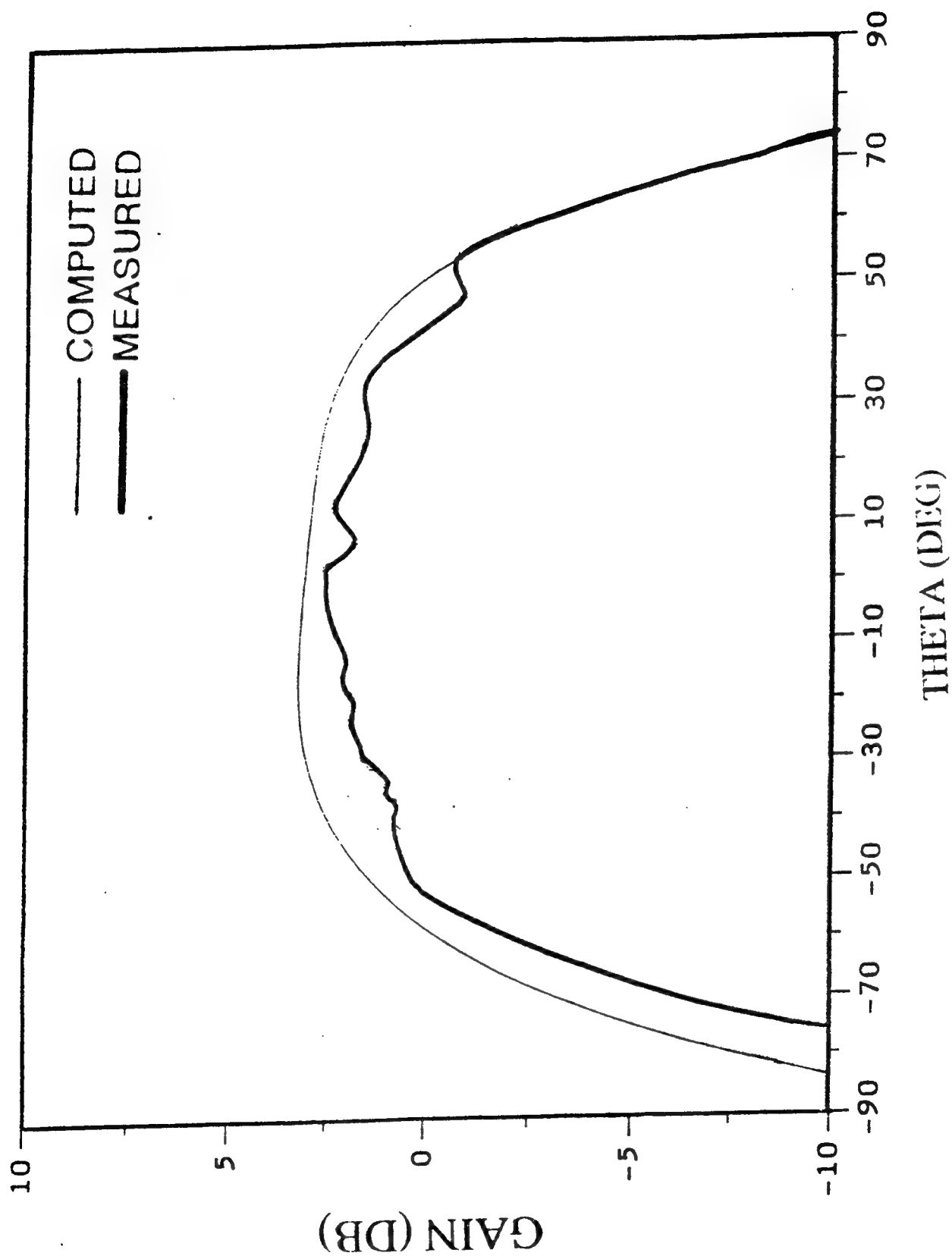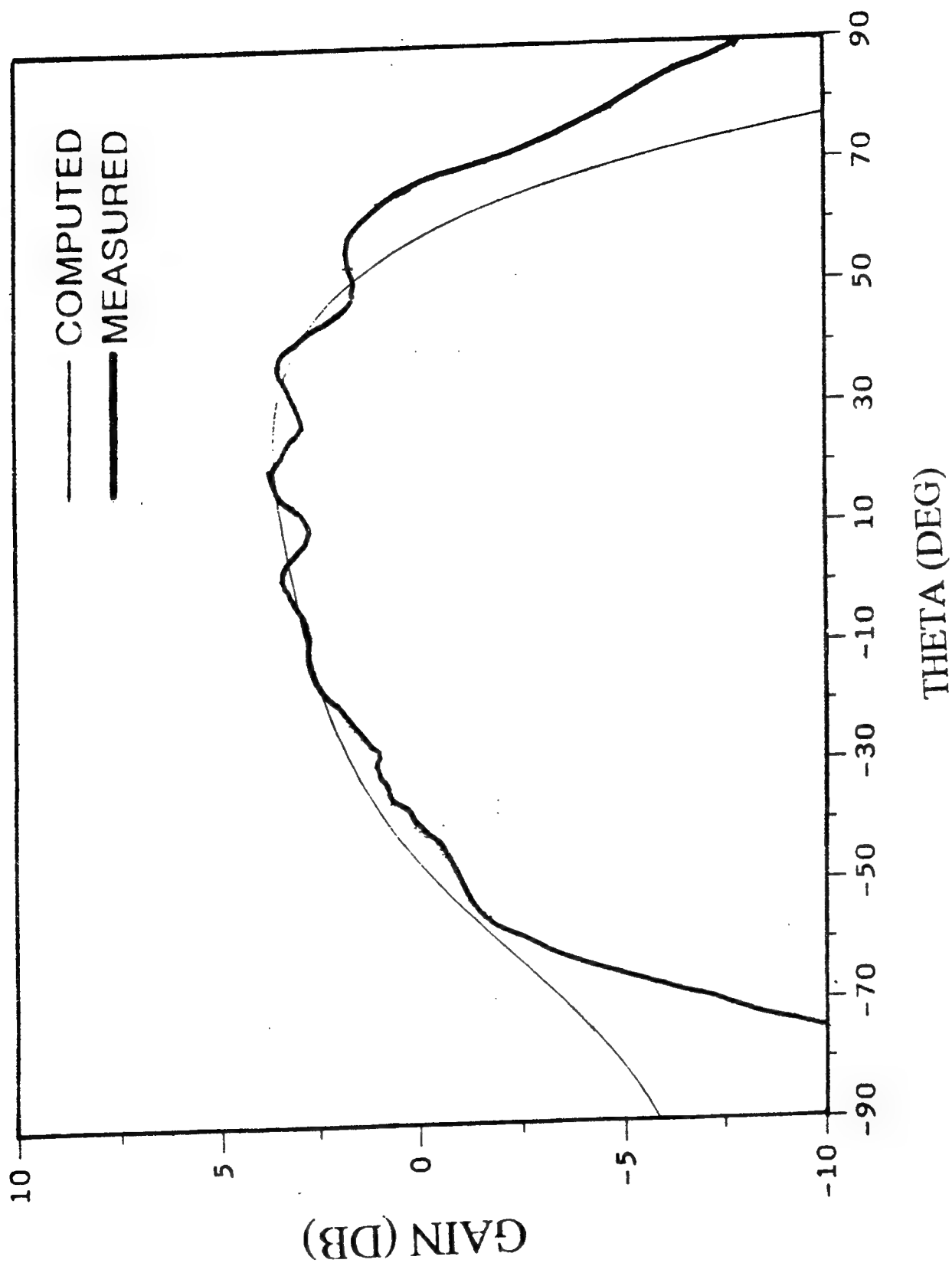
FIGURE 10. CIRCULARLY POLARIZED PATTERNS IN θ - PLANE (φ = 90°)

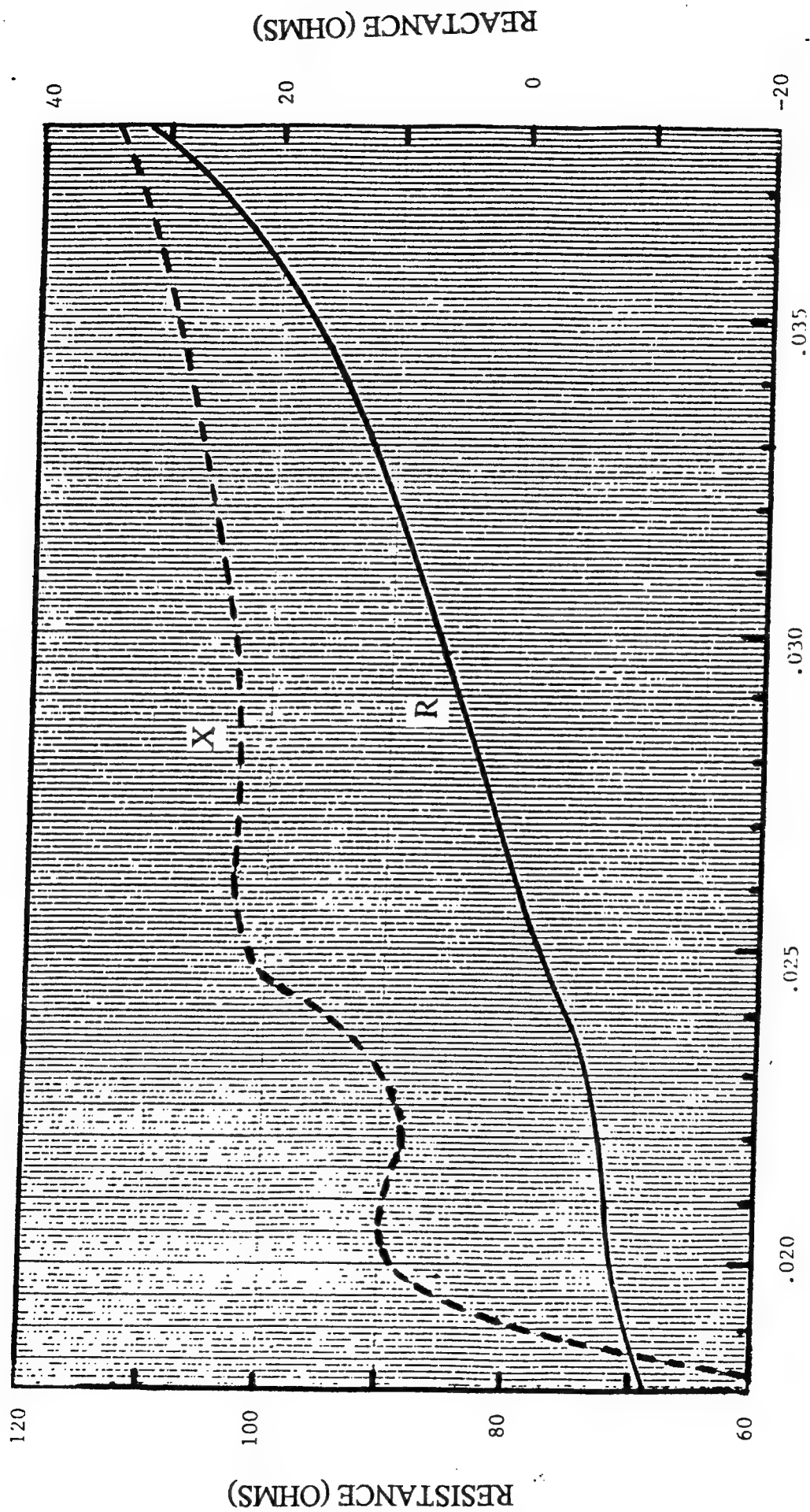FIGURE 11. CIRCULARLY POLARIZED PATTERNS IN θ - PLANE (φ = 135°)

FIGURE 12. INPUT IMPEDANCE DEPENDENCE ON SEGMENT LENGTH

EXPERIMENTAL STUDY ON STACKED RING COUPLED

TRIANGULAR MICROSTRIP ANTENNA

Anjali Naik, P.S. Bhatnagar
Central Electronics Engineering Research Institute
Pilani (Rajasthan) 333031, INDIA

Abstract: A novel configuration for increasing the impedance band width and directivity of the triangular microstrip antenna is reported. The proposed antenna exhibits a 2:1 VSWR bandwidth of 5.2% and gain of about 5.5 dB at the center frequency of 4.3 GHz.

Introduction

In recent years the use and interest in microstrip antennas has become wide spread due to its well known properties. Need for improving the band width of microstrip antenna is well known and several suggestions for triangular structure have been reported by the author [1-3]. A novel configuration is proposed, namely Ring Coupled Triangular Microstrip Antenna (RCTMA), in which triangular patch is gap coupled with triangular ring and one of the non radiating edges is directly coupled through a short section of microstrip line to the triangular ring. Coupling is controlled by the length, width and location of the microstrip section to obtain optimum bandwidth. In addition to bandwidth, to enhance the gain of the proposed antenna, a stacked structure has been incorporated. A triangular patch with same dimensions as RCTMA is fabricated on the similar dielectric substrate (h = 1.6 mm, Er = 2.55) and is space coupled (stacked) only throuh a fringing field and aligned exactly above the lower patch (RCTMA) at optimized

497

height. The lower patch is fed throuh a coaxial probe at suitable position as shown in Fig. 1.

Experimental Results

Initially bandwidth of RCTMA is optimized by bringing the optimum size loop of the impedance locus inside the VSWR = 2 circle using various combinations of design parameters such as gap width, ring width. A short section of the microstrip coupling line is also used to control the input impedance over a wide range of values. A band width of about 210 MHz i.e. 4.92% at the center frequency of 4.3 GHz is obtained with experimentally optimized dimensions of RCTMA as b = 5.1 cm, a = 3 cm, d = 0.5 cm, s = 0.07 cm, 1 = 0.23 cm and fd = 0.45 cm as shown in Fig. 1. RCTMA has been modified by incorporating a stacked structure. The proposed stacked structure offers bandwidth of 5.2% at the center frequency of 4.3 GHZ. The measured gain is about 5.5 dB, which is about 2 times the gain of unstacked RCTMA. The cross polar level is better than 15 dB in the E-plane and 16 dB in H plane in the frequency range between 4.16 GHz and 4.40 GHz.

References

[1] P.S. Bhatnagar, C. Terret et al "Experimental study on broad band triangular microstrip array antenna", Proc. ISAP 89, pp 29-32, Aug. 1989.

[2] P.S. Bhatnagar, C. Terret et al, "Experimental study on stacked triangular microstrip antenna", Electronics Letters, Vol. 22, NR 16, July 86.

[3] P.S. Bhatnagar, C. Terret et al, "Multiple stacked triangular microstrip antenna", IEE International Conference on Antennas and Propagation, UK, 1987.

$b = 5.1\,CM$
$h_1, h_2 = 0.16\,cm$
$t = 2.55\,cm$
G.P.ht= 0.3 cm

Dielectric Substrate

h_2
t
h_1

GROUND PLANE

COAXIAL FEED

$b = 5.1\,cm$
$a = 3\,cm$
$d = 0.5\,cm$
$s = 0.07\,cm$
$l = 0.23\,cm$
fd= 0.45 cm
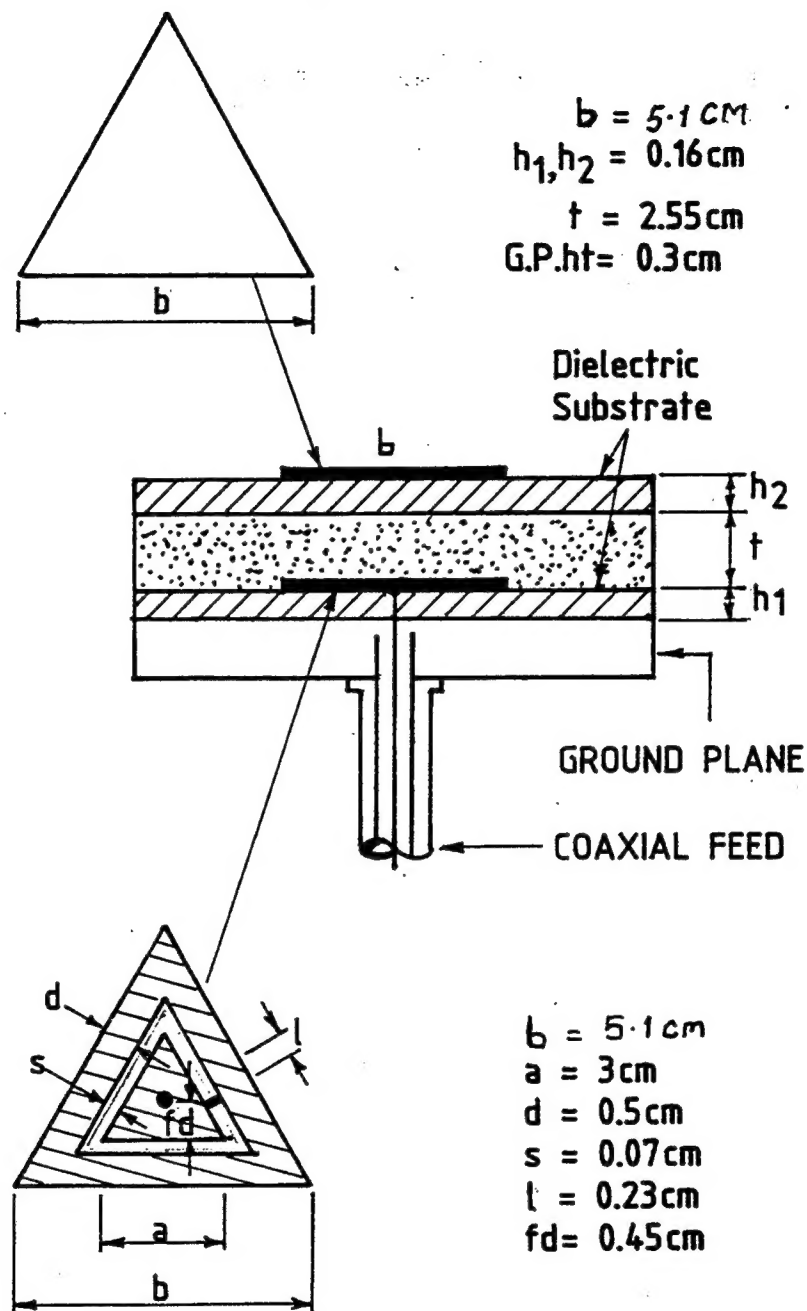
FIG:1 BASIC CONFIGURATION OF STACKED RING
COUPLED TRANGULAR MICROSTRIP ANTENNA.

MARKER 1 ϕ = 4.25 GHz

FIG. 2 INPUT IMPEDANCE CHARACTERISTICS

S$_{11}$
REF 1.0
3   1.0
∇ −0.2068'

C MARKER 3−2
   223.0 MHz
D

VSWR

4.0   4.25   4.5

FREQUENCY (GHz)

FIG. 3 VSWR Vs FREQUENCY VARIATION

MISSION

OF

ROME LABORATORY

Rome Laboratory plans and executes an interdisciplinary program in research, development, test, and technology transition in support of Air Force Command, Control, Communications and Intelligence (C^3I) activities for all Air Force platforms. It also executes selected acquisition programs in several areas of expertise. Technical and engineering support within areas of competence is provided to ESD Program Offices (POs) and other ESD elements to perform effective acquisition of C^3I systems. In addition, Rome Laboratory's technology supports other AFSC Product Divisions, the Air Force user community, and other DOD and non-DOD agencies. Rome Laboratory maintains technical competence and research programs in areas including, but not limited to, communications, command and control, battle management, intelligence information processing, computational sciences and software producibility, wide area surveillance/sensors, signal processing, solid state sciences, photonics, electromagnetic technology, superconductivity, and electronic reliability/maintainability and testability.